# Kongeriget Danmark

Patent application No.:     PA 2002 00532

Date of filing:     10 April 2002

Applicant:     NOVOZYMES A/S
(Name and address)     Krogshøjvej 36
2880 Bagsværd
Denmark

Title: Improved Bacillus Host Cell
IPC: -
The attached documents are exact copies of the filed application

**PRIORITY
DOCUMENT**
SUBMITTED OR TRANSMITTED IN
COMPLIANCE WITH RULE 17.1(a) OR (b)

**Patent- og Varemærkestyrelsen**
Økonomi- og Erhvervsministeriet

11 April 2003

Maiken Lind

**PATENT- OG VAREMÆRKESTYRELSEN**

1

TITLE: Improved Bacillus Host Cell

## TECHNICAL FIELD

*Bacillus sp.* are attractive hosts for the production of heterologous proteins due their ability to

5    secrete proteins directly into the culture medium. They have a high capacity for protein secretion, are genetically highly amenable, nonpathogenic and free of endotoxins, and consequently a large variety of proteins from different organisms have been efficiently produced and secreted in *Bacillus sp. i.e.* in *Bacillus licheniformis.*

In the highly competitive biotech industry, even slightly improved *Bacillus* host cells are in

10    demand, which may provide more attractive production systems, or may even just be alternative production systems.

## BACKGROUND

Many industrial products of commercial interest can be produced biologically in *Bacillus sp.*

15    host cells e.g. heterologous polypeptides, amino acids, carbohydrates etc. Some of these products are sold as process aids, intermediates, or even end-products in the food and feed industries as well as in the pharmaceutical industry. There are increasingly strict regulations that must be complied with when producing such products in microbial production hosts for sale in these industries, for instance traces of antibiotics in the products is seen as a

20    problem. When producing in *Bacillus licheniformis* it is thus desirable to ensure that the host cell is not capable of producing antibiotic compounds native to the cell such as lichenysin, subtilisin, and surfactin.

## SUMMARY

25    A problem to be solved by the present invention is how to obtain a *Bacillus licheniformis* host cell incapable of producing native antibiotic compounds, or how to impair the production of these compounds in said cell. The present invention provides a solution to the problem by providing a *Bacillus licheniformis* host cell which has a reduced capacity to produce one or more polypeptide(s) involved in antibiotic synthesis.

30    Accordingly, in a first aspect the invention relates to a *Bacillus licheniformis* mutant host cell derived from a parent *B. licheniformis* host cell, which mutant host cell is mutated in one or more gene(s) encoding one or more polypeptide(s) involved in antibiotic synthesis which is at least 80% identical to one or more of the polypeptides shown in SEQ ID NO's: 2 to 10, preferably at least 85% identical, more preferably at least 90% identical, still more preferably

35    at least 95% identical, and most preferably at least 97% identical to one or more of the polypeptides shown in SEQ ID NO's: 2 to 10, wherein the mutant host cell expresses at least

5% less of the one or more polypeptide(s) involved in antibiotic synthesis than the parent host cell, when they are cultivated under comparable conditions.

Preferably the mutant host cell expresses at least 10% less, more preferably at least 20% less, still more preferably at least 30% less, even more preferably at least 40% less, yet more preferably at least 50% less, or at least 60% less, or at least 70% less, or at least 80%, or most preferably at least 90% less of the one or more polypeptide(s) involved in antibiotic synthesis than the parent host cell, when they are cultivated under comparable conditions. Most preferably the mutant host cell expresses absolutely nothing of the one or more polypeptide(s) involved in antibiotic synthesis.

Comparable conditions of cultivation must be used in order to compare the expression level of the one or more polypeptide(s) involved in antibiotic synthesis in a mutant host cell of the invention with that in a parent host cell. They are cultivated separately under identical conditions in identical setups, of course allowing for the usual standard deviations of the operating parameters normally associated with growth experiments, such as temperature control etc. The quantification of the expression level of the one or more polypeptide(s) is done by standard text-book assay techniques as known in the art e.g. mRNA quantification or immuno-based assays.

In a second aspect the invention relates to a process for producing at least one product of interest in a *Bacillus licheniformis* mutant host cell, comprising cultivating a *B.licheniformis* mutant host cell as defined in the previous aspect in a suitable medium, whereby the said product is produced.

Finally, an aspect of the invention relates to a use of a *Bacillus licheniformis* mutant host cell as defined in the first aspect for producing at least one product of interest comprising cultivating the mutant host cell in a suitable medium whereby the said product is produced.

**DEFINITIONS**

Nucleic acid construct: When used herein, the term "nucleic acid construct" means a nucleic acid molecule, either single- or double-stranded, which is isolated from a naturally occurring gene or which has been modified to contain segments of nucleic acids in a manner that would not otherwise exist in nature. The term nucleic acid construct is synonymous with the term "expression cassette" when the nucleic acid construct contains the control sequences required for expression of a coding sequence of the present invention.

Control sequence: The term "control sequences" is defined herein to include all components, which are necessary or advantageous for the expression of a polypeptide of the present invention. Each control sequence may be native or foreign to the nucleotide sequence encoding the polypeptide. Such control sequences include, but are not limited to, a leader,

5    polyadenylation sequence, propeptide sequence, promoter, signal peptide sequence, and transcription terminator. At a minimum, the control sequences include a promoter, and transcriptional and translational stop signals. The control sequences may be provided with linkers for the purpose of introducing specific restriction sites facilitating ligation of the control sequences with the coding region of the nucleotide sequence encoding a polypeptide.

10

Operably linked: The term "operably linked" is defined herein as a configuration in which a control sequence is appropriately placed at a position relative to the coding sequence of the DNA sequence such that the control sequence directs the expression of a polypeptide.

15    Coding sequence: When used herein the term "coding sequence" is intended to cover a nucleotide sequence, which directly specifies the amino acid sequence of its protein product. The boundaries of the coding sequence are generally determined by an open reading frame, which usually begins with the ATG start codon. The coding sequence typically include DNA, cDNA, and recombinant nucleotide sequences.

20

Expression: In the present context, the term "expression" includes any step involved in the production of the polypeptide including, but not limited to, transcription, post-transcriptional modification, translation, post-translational modification, and secretion.

25    Expression vector: In the present context, the term "expression vector" covers a DNA molecule, linear or circular, that comprises a segment encoding a polypeptide of the invention, and which is operably linked to additional segments that provide for its transcription.

30    **DETAILED DISCLOSURE**

A *Bacillus licheniformis* mutant host cell derived from a parent *B. licheniformis* host cell, which mutant host cell is mutated in one or more gene(s) encoding one or more polypeptide(s) involved in antibiotic synthesis which is at least 80% identical to one or more of the polypeptides shown in SEQ ID NO's: 2 to 10, wherein the mutant host cell expresses

35    at least 5% less of the one or more polypeptide(s) involved in antibiotic synthesis than the parent host cell, when they are cultivated under comparable conditions.

4

The term "parent host cell" in the context of the present invention means a cell which is genetically identical, or isogenic, to the progeny mutant or mutant cell of the present invention, except for the mutated one or more gene(s) encoding one or more polypeptide(s) involved in antibiotic synthesis in said mutant.

5

The degree of identity, or %-identity of polypeptide sequences can suitably be investigated by aligning the sequences using a computer program known in the art, such as "GAP" provided in the GCG program package (Program Manual for the Wisconsin Package, Version 8, August 1994, Genetics Computer Group, 575 Science Drive, Madison, Wisconsin,

10    USA 53711)(Needleman, S.B. and Wunsch, C.D., (1970), Journal of Molecular Biology, 48, 443-453). Using GAP with the following settings for DNA sequence comparison: GAP creation penalty of 5.0 and GAP extension penalty of 0.3".

An object of the present invention is to provide a culture medium free of antibiotics so as to

15    be able to reduce the product purification to a minimum, and to comply with regulatory requirements. This may be done according to the invention by reducing or even completely abolishing the expression of one or more gene(s) encoding a native polypeptide(s) involved in antibiotic synthesis via mutagenisation of that (those) gene(s). One of the very well-known method of ensuring that a gene is not expressed into an active polypeptide within a cell is

20    simply to delete or partially delete the encoding gene. Many techniques have been described in the art on how to specifically delete or partially delete one or more gene(s) in the genome of a cell, and certainly from the genome of a *Bacillus licheniformis* cell (see e.g. Novozymes A/S WO 01/90393, Novozymes A/S WO 02/00907). Accordingly, a preferred embodiment of the present invention relates to a host cell of the first aspect, which is mutated by a partial or

25    complete deletion of the one or more gene(s) encoding the one or more polypeptide(s) involved in antibiotic synthesis.

A preferred embodiment of the present invention relates to a host cell of the first aspect, which is mutated in two or more genes encoding two or more polypeptides involved in

30    antibiotic synthesis.

The product of interest to be produced by the mutant host cell of the first aspect may be one or more polypeptide(s) encoded by one or more heterologous gene(s). Consequently, a preferred embodiment of the present invention relates to a host cell of the first aspect, which

35    comprises one or more heterologous gene(s) encoding one or more heterologous polypeptide(s).

In the industrial production of polypeptides it is of interest to achieve a product yield as high as possible. One way to increase the yield is to increase the copy number of a gene encoding a polypeptide of interest. This can be done by placing the gene on a high copy number plasmid. However, plasmids are unstable and are often lost from the host cells if

5 there is no selective pressure during the cultivation of the host cells. Another way to increase the copy number of the gene of interest is to integrate it into the host cell chromosome in multiple copies. Integration of two genes has been described in WO 91/09129 and WO 94/14968 (Novozymes A/S) the content of which is hereby incorporated by reference. A preferred embodiment of the present invention relates to a host cell of the first aspect,

10 wherein the heterologous gene(s) is present in at least two copies, preferably at least 4 copies, and most preferably at least 6 copies. In another embodiment the heterologous gene(s) is present in at least ten copies. If carried on a plasmid the gene(s) may be present in several hundred copies per cell, so in a still further embodiment of the present invention the heterologous gene(s) is present in at least 100 copies.

15

Integration of two genes closely spaced in anti-parallel tandem to achieve better stability has been described in WO 99/41358 (Novozymes A/S) the content of which is hereby incorporated by reference, as well as the stable chromosomal multi-copy integration of genes described in WO 02/00907 (Novozymes A/S) the content of which is incorporated herein by

20 reference. A preferred embodiment of the present invention relates to a host cell of the first aspect, wherein the heterologous gene(s) are stably integrated into the genome of the cell.

Selection of chromosomal integrant has for convenience resulted in the use of selectable markers such as antibiotic resistance markers. However it is desirable if possible to avoid the

25 use of antibiotic marker genes. WO 01/90393 discloses a method for the integration of a gene in the chromosome of a host cell without leaving antibiotic resistance markers behind in the strain, the content of which is hereby incorporated by reference A preferred embodiment of the present invention relates to a host cell of the first aspect wherein the heterologous gene(s) is integrated into the genome of the cell without leaving any antibiotic resistance

30 marker gene(s) at the site of integration.

The present invention also relates to nucleic acid constructs comprising a nucleotide sequence encoding a product of interest, which may be operably linked to one or more control sequences that direct the expression of the coding sequence in a suitable host cell

35 under conditions compatible with the control sequences.

A nucleotide sequence encoding a polypeptide of interest may be manipulated in a variety of ways to provide for expression of the polypeptide. Manipulation of the nucleotide sequence prior to its insertion into a vector may be desirable or necessary depending on the expression vector. The techniques for modifying nucleotide sequences utilizing recombinant DNA

5 methods are well known in the art.

Other ways of increasing the product yield would be to increase promoter activity of the specific promoter regulating the expression of a specific gene of interest. Also a more general increase in the activity of several promoters at the same time could lead to an

10 improved product yield. The control sequence may be an appropriate promoter sequence, a nucleotide sequence which is recognized by a host cell for expression of the nucleotide sequence. The promoter sequence contains transcriptional control sequences, which mediate the expression of the polypeptide. The promoter may be any nucleotide sequence which shows transcriptional activity in the host cell of choice including mutant, truncated, and

15 hybrid promoters, and may be obtained from genes encoding extracellular or intracellular polypeptides either homologous or heterologous to the host cell.

Examples of suitable promoters for directing the transcription of the nucleic acid constructs of the present invention, especially in a bacterial host cell, are the promoters obtained from the

20 E. coli lac operon, Streptomyces coelicolor agarase gene (dagA), Bacillus subtilis levansucrase gene (sacB), Bacillus licheniformis alpha-amylase gene (amyL), Bacillus stearothermophilus maltogenic amylase gene (amyM), Bacillus amyloliquefaciens alpha-amylase gene (amyQ), Bacillus licheniformis penicillinase gene (penP), Bacillus subtilis xylA and xylB genes, and prokaryotic beta-lactamase gene (Villa-Kamaroff et al., 1978,

25 Proceedings of the National Academy of Sciences USA 75: 3727-3731), as well as the tac promoter (DeBoer et al., 1983, Proceedings of the National Academy of Sciences USA 80: 21-25). Further promoters are described in "Useful proteins from recombinant bacteria" in Scientific American, 1980, 242: 74-94; and in Sambrook et al., 1989, supra.

30 Other useful promoters are described in WO 93/10249, WO 98/07846, and WO 99/43835 (Novozymes A/S) the contents of which are incorporated fully herein by reference. A preferred embodiment of the present invention relates to a host cell of the first aspect, wherein the heterologous gene(s) are transcribed from a heterologous promoter or from an artificial promoter.

35

The control sequence may also be a suitable transcription terminator sequence, a sequence recognized by a host cell to terminate transcription. The terminator sequence is operably

linked to the 3' terminus of the nucleotide sequence encoding the polypeptide. Any terminator which is functional in the host cell of choice may be used in the present invention.

The control sequence may also be a suitable leader sequence, a nontranslated region of an
5    mRNA which is important for translation by the host cell. The leader sequence is operably linked to the 5' terminus of the nucleotide sequence encoding the polypeptide. Any leader sequence that is functional in the host cell of choice may be used in the present invention.

The control sequence may also be a polyadenylation sequence, a sequence operably linked
10   to the 3' terminus of the nucleotide sequence and which, when transcribed, is recognized by the host cell as a signal to add polyadenosine residues to transcribed mRNA. Any polyadenylation sequence which is functional in the host cell of choice may be used in the present invention.

15   The control sequence may also be a signal peptide coding region that codes for an amino acid sequence linked to the amino terminus of a polypeptide and directs the encoded polypeptide into the cell's secretory pathway. The 5' end of the coding sequence of the nucleotide sequence may inherently contain a signal peptide coding region naturally linked in translation reading frame with the segment of the coding region which encodes the secreted
20   polypeptide. Alternatively, the 5' end of the coding sequence may contain a signal peptide coding region which is foreign to the coding sequence. The foreign signal peptide coding region may be required where the coding sequence does not naturally contain a signal peptide coding region. Alternatively, the foreign signal peptide coding region may simply replace the natural signal peptide coding region in order to enhance secretion of the
25   polypeptide. However, any signal peptide coding region which directs the expressed polypeptide into the secretory pathway of a host cell of choice may be used in the present invention.

Effective signal peptide coding regions for bacterial host cells are the signal peptide coding
30   regions obtained from the genes for Bacillus NCIB 11837 maltogenic amylase, Bacillus stearothermophilus alpha-amylase, Bacillus licheniformis subtilisin, Bacillus licheniformis beta-lactamase, Bacillus stearothermophilus neutral proteases (nprT, nprS, nprM), and Bacillus subtilis prsA. Further signal peptides are described by Simonen and Palva, 1993, Microbiological Reviews 57: 109-137.

35

The control sequence may also be a propeptide coding region that codes for an amino acid sequence positioned at the amino terminus of a polypeptide. The resultant polypeptide is

known as a proenzyme or propolypeptide (or a zymogen in some cases). A propolypeptide is generally inactive and can be converted to a mature active polypeptide by catalytic or autocatalytic cleavage of the propeptide from the propolypeptide. The propeptide coding region may be obtained from the genes for Bacillus subtilis alkaline protease (aprE), Bacillus subtilis neutral protease (nprT), Saccharomyces cerevisiae alpha-factor, Rhizomucor miehei aspartic proteinase, and Myceliophthora thermophila laccase (WO 95/33836).

Where both signal peptide and propeptide regions are present at the amino terminus of a polypeptide, the propeptide region is positioned next to the amino terminus of a polypeptide and the signal peptide region is positioned next to the amino terminus of the propeptide region.

It may also be desirable to add regulatory sequences which allow the regulation of the expression of the polypeptide relative to the growth of the host cell. Examples of regulatory systems are those which cause the expression of the gene to be turned on or off in response to a chemical or physical stimulus, including the presence of a regulatory compound. Regulatory systems in prokaryotic systems include the lac, tac, and trp operator systems. In yeast, the ADH2 system or GAL1 system may be used. In eukaryotic systems, these include the dihydrofolate reductase gene which is amplified in the presence of methotrexate, and the metallothionein genes which are amplified with heavy metals. In these cases, the nucleotide sequence encoding the polypeptide would be operably linked with the regulatory sequence.

The present invention also relates to recombinant expression vectors comprising the nucleic acid construct of the invention. The various nucleotide and control sequences described above may be joined together to produce a recombinant expression vector which may include one or more convenient restriction sites to allow for insertion or substitution of the nucleotide sequence encoding the polypeptide at such sites. Alternatively, the nucleotide sequence of the present invention may be expressed by inserting the nucleotide sequence or a nucleic acid construct comprising the sequence into an appropriate vector for expression. In creating the expression vector, the coding sequence is located in the vector so that the coding sequence is operably linked with the appropriate control sequences for expression.

The recombinant expression vector may be any vector (e.g., a plasmid or virus) which can be conveniently subjected to recombinant DNA procedures and can bring about the expression of the nucleotide sequence. The choice of the vector will typically depend on the compatibility of the vector with the host cell into which the vector is to be introduced. The vectors may be linear or closed circular plasmids.

The vector may be an autonomously replicating vector, i.e., a vector which exists as an extrachromosomal entity, the replication of which is independent of chromosomal replication, e.g., a plasmid, an extrachromosomal element, a minichromosome, or an artificial

5    chromosome.

The vector may contain any means for assuring self-replication. Alternatively, the vector may be one which, when introduced into the host cell, is integrated into the genome and replicated together with the chromosome(s) into which it has been integrated. Furthermore,

10   a single vector or plasmid or two or more vectors or plasmids which together contain the total DNA to be introduced into the genome of the host cell, or a transposon may be used.

The vectors of the present invention preferably contain one or more selectable markers which permit easy selection of transformed cells. A selectable marker is a gene the product

15   of which provides for biocide or viral resistance, resistance to heavy metals, prototrophy to auxotrophs, and the like.

Examples of bacterial selectable markers are the dal genes from Bacillus subtilis or Bacillus licheniformis, or markers which confer antibiotic resistance such as ampicillin, kanamycin,

20   chloramphenicol or tetracycline resistance.

The vectors of the present invention preferably contain an element(s) that permits stable integration of the vector into the host cell's genome or autonomous replication of the vector in the cell independent of the genome.

25

For integration into the host cell genome, the vector may rely on the nucleotide sequence encoding the polypeptide or any other element of the vector for stable integration of the vector into the genome by homologous or nonhomologous recombination. Alternatively, the vector may contain additional nucleotide sequences for directing integration by homologous

30   recombination into the genome of the host cell. The additional nucleotide sequences enable the vector to be integrated into the host cell genome at a precise location(s) in the chromosome(s).   To increase the likelihood of integration at a precise location, the integrational elements should preferably contain a sufficient number of nucleotides, such as 100 to 1,500 base pairs, preferably 400 to 1,500 base pairs, and most preferably 800 to

35   1,500 base pairs, which are highly homologous with the corresponding target sequence to enhance the probability of homologous recombination. The integrational elements may be any sequence that is homologous with the target sequence in the genome of the host cell.

Furthermore, the integrational elements may be non-encoding or encoding nucleotide sequences. On the other hand, the vector may be integrated into the genome of the host cell by non-homologous recombination.

5    For autonomous replication, the vector may further comprise an origin of replication enabling the vector to replicate autonomously in the host cell in question. Examples of bacterial origins of replication are the origins of replication of plasmids pBR322, pUC19, pACYC177, and pACYC184 permitting replication in E. coli, and pUB110, pE194, pTA1060, and pAMß1 permitting replication in Bacillus. The origin of replication may be one having a mutation
10   which makes its functioning temperature-sensitive in the host cell (see, e.g., Ehrlich, 1978, Proceedings of the National Academy of Sciences USA 75: 1433).

More than one copy of a nucleotide sequence of the present invention may be inserted into the host cell to increase production of the gene product. An increase in the copy number of
15   the nucleotide sequence can be obtained by integrating at least one additional copy of the sequence into the host cell genome or by including an amplifiable selectable marker gene with the nucleotide sequence where cells containing amplified copies of the selectable marker gene, and thereby additional copies of the nucleotide sequence, can be selected for by cultivating the cells in the presence of the appropriate selectable agent.
20
The procedures used to ligate the elements described above to construct the recombinant expression vectors of the present invention are well known to one skilled in the art (see, e.g., Sambrook et al., 1989, supra).

25   The introduction of a vector into a bacterial host cell may, for instance, be effected by protoplast transformation (see, e.g., Chang and Cohen, 1979, Molecular General Genetics 168: 111-115), using competent cells (see, e.g., Young and Spizizin, 1961, Journal of Bacteriology 81: 823-829, or Dubnau and Davidoff-Abelson, 1971, Journal of Molecular Biology 56: 209-221), electroporation (see, e.g., Shigekawa and Dower, 1988, Biotechniques
30   6: 742-751), or conjugation (see, e.g., Koehler and Thorne, 1987, Journal of Bacteriology 169: 5771-5278).

A preferred embodiment of the present invention relates to a host cell of the first aspect, wherein the heterologous gene(s) are comprised in an operon, preferably a polycistronic
35   operon. The term "operon" in the context of the present invention means a polynucleotide comprising several genes that are clustered and perhaps even transcribed together into a polycistronic mRNA, e.g. genes coding for the enzymes of a metabolic pathway. The

transcription of an operon may be initiated at a promoter region and controlled by a neighboring regulatory gene, which encodes a regulatory protein, which in turn binds to the operator sequence in the operon to respectively inhibit or enhance the transcription. The gene or the operon can be carried on a suitable plasmid that can be stably maintained, e.g.

5    capable of stable autonomous replication in the host cell (the choice of plasmid will typically depend on the compatibility of the plasmid with the host cell into which the plasmid is to be introduced) or it can be carried on the chromosome of the host. The said gene may be endogenous to the host cell in which case the product of interest is a protein naturally produced by the host cell and in most cases the gene will be in it normal position on the

10    chromosome. If the gene encoding the product of interest is an exogenous gene, the gene could either be carried on a suitable plasmid or it could be integrated on the host chromosome. In one embodiment of the invention the eubacterium is a recombinant eubacterium. Also the product of interest may in another embodiment be a recombinant protein.

15

The product of interest is any gene product or product of a metabolic pathway which is industrially useful and which can be produced in a bacterial cell such as a *B. licheniformis.*

In one preferred embodiment, the heterologous polypeptide(s) is an antimicrobial peptide, or

20    a fusion peptide comprising a peptide part which in its native form has antimicrobial activity.

In another preferred embodiment, the heterologous polypeptide(s) has biosynthetic activity and produces a compound or an intermediate of interest.

25    Yet another embodiment relates to a host cell of the first aspect, wherein the compound or intermediate of interest comprises vitamins, amino acids, antibiotics, carbohydrates, or surfactants, and preferably the carbohydrates comprise hyaluronic acid.

In one embodiment the heterologous polypeptide(s) is an enzyme, particularly the enzyme is

30    an enzyme of a class selected from the group of enzyme classes consisting of oxidoreductases (EC 1), transferases (EC 2), hydrolases (EC 3), lyases (EC 4), isomerases (EC 5), and ligases (EC 6). Preferably the enzyme is an enzyme with an activity selected from the group consisting of aminopeptidase, amylase, amyloglucosidase, mannanase, carbohydrase, carboxypeptidase, catalase, cellulase, chitinase, cutinase, cyclodextrin

35    glycosyltransferase, deoxyribonuclease, esterase, galactosidase, beta-galactosidase, glucoamylase, glucose oxidase, glucosidase, haloperoxidase, hemicellulase, invertase, isomerase, laccase, ligase, lipase, lyase, mannosidase, oxidase, pectinase, peroxidase,

phytase, phenoloxidase, polyphenoloxidase, protease, ribonuclease, transferase, transglutaminase, or xylanase. Preferably the enzyme is an amylase or a mannanase.

A second aspect of the invention relates to a process for producing at least one product of interest in a *Bacillus licheniformis* mutant host cell, comprising cultivating a *B.licheniformis* mutant host cell as defined in the first aspect of the invention in a suitable medium, whereby the said product is produced. One embodiment relates to a process of the second aspect, further comprising isolating or purifying the product of interest. Suitable media for the cultivation is described below as well as methods for the purification or isolation of the produced product which is an optional additional step to the process of the present invention.

In the production methods of the present invention, the cells are cultivated in a nutrient medium suitable for production of the polypeptide using methods known in the art. For example, the cell may be cultivated by shake flask cultivation, small-scale or large-scale fermentation (including continuous, batch, fed-batch, or solid state fermentations) in laboratory or industrial fermentors performed in a suitable medium and under conditions allowing the polypeptide to be expressed and/or isolated. The cultivation takes place in a suitable nutrient medium comprising carbon and nitrogen sources and inorganic salts, using procedures known in the art. Suitable media are available from commercial suppliers or may be prepared according to published compositions (*e.g.*, in catalogues of the American Type Culture Collection). If the polypeptide is secreted into the nutrient medium, the polypeptide can be recovered directly from the medium. If the polypeptide is not secreted, it can be recovered from cell lysates.

The medium used to culture the cells may be any conventional medium suitable for growing the host cells, such as minimal or complex media containing appropriate supplements. Suitable media are available from commercial suppliers or may be prepared according to published recipes (e.g. in catalogues of the American Type Culture Collection). The media are prepared using procedures known in the art (see, *e.g.*, references for bacteria and yeast; Bennett, J.W. and LaSure, L., editors, *More Gene Manipulations in Fungi*, Academic Press, CA, 1991).

The polypeptides may be detected using methods known in the art that are specific for the polypeptides. These detection methods may include use of specific antibodies, formation of an enzyme product, or disappearance of an enzyme substrate. For example, an enzyme assay may be used to determine the activity of the polypeptide as described herein.

The resulting polypeptide may be recovered by methods known in the art. For example, the polypeptide may be recovered from the nutrient medium by conventional procedures including, but not limited to, centrifugation, filtration, extraction, spray-drying, evaporation, or precipitation.

5

The polypeptides of the present invention may be purified by a variety of procedures known in the art including, but not limited to, chromatography (*e.g.*, ion exchange, affinity, hydrophobic, chromatofocusing, and size exclusion), electrophoretic procedures (*e.g.*, preparative isoelectric focusing), differential solubility (*e.g.*, ammonium sulfate precipitation),

10    SDS-PAGE, or extraction (see, *e.g.*, *Protein Purification*, J.-C. Janson and Lars Ryden, editors, VCH Publishers, New York, 1989).

A third aspect of the present invention relates to the use of a *Bacillus licheniformis* mutant host cell as defined in the first aspect for producing at least one product of interest

15    comprising cultivating the mutant host cell in a suitable medium whereby the said product is produced, and optionally isolating or purifying the produced product.

20

## CLAIMS

1. A *Bacillus licheniformis* mutant host cell derived from a parent *B. licheniformis* host cell, which mutant host cell is mutated in one or more gene(s) encoding one or more polypeptide(s) involved in antibiotic synthesis which is at least 80% identical to one or more
5    of the polypeptides shown in SEQ ID NO's: 2 to 10, wherein the mutant host cell expresses at least 5% less of the one or more polypeptide(s) involved in antibiotic synthesis than the parent host cell, when they are cultivated under comparable conditions.

2. The host cell according to claim 1, which is mutated by a partial or complete deletion of
10    the one or more gene(s) encoding the one or more polypeptide(s) involved in antibiotic synthesis.

3. The host cell according to any of claims 1 – 2, which is mutated in two or more genes encoding two or more polypeptides involved in antibiotic synthesis.
15

5. The host cell according to any of claims 1 – 4, which comprises one or more heterologous gene(s) encoding one or more heterologous polypeptide(s).

6. The host cell according to claim 5, wherein the heterologous gene(s) is present in at least
20    two copies.

7. The host cell according to claim 5 or 6, wherein the heterologous gene(s) are stably integrated into the genome of the cell.

25    8. The host cell according to any of claims 5 - 7, wherein the heterologous gene(s) is integrated into the genome of the cell without leaving any antibiotic resistance marker genes at the site of integration.

9. The host cell according to any of claims 5 - 8, wherein the heterologous gene(s) are
30    transcribed from a heterologous promoter or from an artificial promoter.

10. The host cell according to any of claim 5 – 9, wherein the heterologous gene(s) are comprised in an operon, preferably a polycistronic operon.

35    11. The host cell according to any of claims 5 – 10, wherein the heterologous polypeptide(s) is an antimicrobial peptide, or a fusion peptide comprising a peptide part which in its native form has antimicrobial activity.

12. The host cell according to any of claims 5 – 10, wherein the heterologous polypeptide(s) has biosynthetic activity and produces a compound or an intermediate of interest.

13. The host cell according to claim 12, wherein the compound or intermediate of interest comprises vitamins, amino acids, antibiotics, carbohydrates, or surfactants.

14. The host cell according to claim 13, wherein the carbohydrates comprise hyaluronic acid.

15. The host cell according to any of claims 5 – 10, wherein the heterologous polypeptide(s) is an enzyme, preferably a secreted enzyme.

16. The host cell according to claim 15, wherein the enzyme is is an enzyme of a class selected from the group of enzyme classes consisting of oxidoreductases (EC 1), transferases (EC 2), hydrolases (EC 3), lyases (EC 4), isomerases (EC 5), and ligases (EC 6).

17. The host cell according to claim 16, wherein the enzyme is an enzyme with an activity selected from the group of enzyme activities consisting of aminopeptidase, amylase, amyloglucosidase, mannanase, carbohydrase, carboxypeptidase, catalase, cellulase, chitinase, cutinase, cyclodextrin glycosyltransferase, deoxyribonuclease, esterase, galactosidase, beta-galactosidase, glucoamylase, glucose oxidase, glucosidase, haloperoxidase, hemicellulase, invertase, isomerase, laccase, ligase, lipase, lyase, mannosidase, oxidase, pectinase, peroxidase, phytase, phenoloxidase, polyphenoloxidase, protease, ribonuclease, transferase, transglutaminase, and xylanase.

18. The host cell according to claim 17, wherein the enzyme is an amylase or a mannanase.

19. A process for producing at least one product of interest in a *Bacillus licheniformis* mutant host cell, comprising cultivating a *B.licheniformis* mutant host cell as defined in any of the claims 1 - 18 in a suitable medium, whereby the said product is produced.

20. The process according to claim 19, further comprising isolating or purifying the product of interest.

21. A use of a *Bacillus licheniformis* mutant host cell as definde in any of the claims 1 - 18 for producing at least one product of interest comprising cultivating the mutant host cell in a suitable medium whereby the said product is produced.

5   22. The use according to claim 21 further comprising isolating or purifying the product of interest.

**ABSTRACT**

**TITLE:** Improved Bacillus Host Cell.

A *Bacillus licheniformis* mutant host cell derived from a parent *B. licheniformis* host cell, which mutant host cell is mutated in one or more gene(s) encoding one or more

5    polypeptide(s) involved in antibiotic synthesis which is at least 80% identical to one or more of the polypeptides shown in SEQ ID NO's: 2 to 10, wherein the mutant host cell expresses at least 5% less of the one or more polypeptide(s) involved in antibiotic synthesis than the parent host cell, when they are cultivated under comparable conditions.

SEQUENCE LISTING

Patent- og
Varemærkestyrelsen

1 0 APR. 2002
Modtaget

<110>  Novozymes A/S

        Jørgensen, Steen Troels

        Rasmussen, Michael Dolberg

        Andersen, Jens Tønne

        Olesen, Peter Bjarke

        Clausen, Ib Groth


<120>  Improved Bacillus Host Cell


<130>  10297.000


<160>  10


<170>  PatentIn version 3.1


<210>  1
<211>  1602
<212>  DNA
<213>  Bacillus licheniformis


<220>
<221>  CDS
<222>  (501)..(1421)
<223>


<400>  1
aacagccaga aaggagacga agaacgtctc tatttaacag agcttgatat gttctgcgtc      60
catttcgggt caaggatata ttcactatta tgaaaccaat ttttccgatt tccgtataaa     120
caaagcgtgg aacctgttaa aggacaagaa aaatcgacag ggatataaaa atcggcaagc     180
tttccaaagg gaaccggggc aggctgaagc tagcgctcac attggcgagc gatgcagcgg     240
tcgtgctgtt tgatgaacca tccggtctcc gatggtgagg gagtcgacag tcaaaagcct     300
tcttacatat ttggagcttg acaagcagac aatcgccatc gccacctatg aaattgaaca     360
aatattggat gtcgtcttaa atggccgaat tgaagcaaaa aaggatgtcg aacagctgcg     420

Page 1

```
cgaggaatcc ggaatgtcgc ttttggaatg gctgaaacgg tttattaaag aaaaaaattg      480

aaaaagcggg aggagaaaca ttg gaa aca ctt ttg gaa tta aaa aat gta tca      533
                      Leu Glu Thr Leu Leu Glu Leu Lys Asn Val Ser
                      1               5                   10

aaa acg atc agg ggg aaa aag atc atc gag ggc ttg agt ttt gac gtg       581
Lys Thr Ile Arg Gly Lys Lys Ile Ile Glu Gly Leu Ser Phe Asp Val
            15              20                  25

cgg gca ggc gag ata ttc ggc ttc ctg ggg ccg aac ggc gcc gga aaa      629
Arg Ala Gly Glu Ile Phe Gly Phe Leu Gly Pro Asn Gly Ala Gly Lys
            30              35              40

acg acg acg atc cgg atg att gtc gga cat atg agc att acc gcc ggt      677
Thr Thr Thr Ile Arg Met Ile Val Gly His Met Ser Ile Thr Ala Gly
        45              50              55

gaa atc gcc gtg tgc ggc gtc agc gta aaa gaa aac ttt gaa aag gct      725
Glu Ile Ala Val Cys Gly Val Ser Val Lys Glu Asn Phe Glu Lys Ala
60              65              70              75

gca cgg cat atc gga gcg atc gtt gaa aac ccg gag ctt tat aaa ttt      773
Ala Arg His Ile Gly Ala Ile Val Glu Asn Pro Glu Leu Tyr Lys Phe
                80              85              90

tta acg ggc tac caa aac ctt cag caa tac gcg cgc atg acg aaa ggc      821
Leu Thr Gly Tyr Gln Asn Leu Gln Gln Tyr Ala Arg Met Thr Lys Gly
            95              100             105

gtg acg aaa aag aaa att gac gaa atc gtc gag ctc gtc gga ttg aaa      869
Val Thr Lys Lys Lys Ile Asp Glu Ile Val Glu Leu Val Gly Leu Lys
        110             115             120

aac agg atc aac gac aag gtc aaa acg tat tcg tta gga atg aga caa      917
Asn Arg Ile Asn Asp Lys Val Lys Thr Tyr Ser Leu Gly Met Arg Gln
125             130             135

agg ctt ggc ctt gcc caa agc ctt ttg cac gat cca aag ctg ttg att      965
Arg Leu Gly Leu Ala Gln Ser Leu Leu His Asp Pro Lys Leu Leu Ile
140             145             150             155

ctc gat gag ccg acg aac ggg ctt gat ccg gca ggt atc cgg gaa atc      1013
Leu Asp Glu Pro Thr Asn Gly Leu Asp Pro Ala Gly Ile Arg Glu Ile
                160             165             170

cgt gac tat tta aga aag ctg acg aga gaa aag gga atg gcg gtc atc      1061
Arg Asp Tyr Leu Arg Lys Leu Thr Arg Glu Lys Gly Met Ala Val Ile
            175             180             185

gtt tca agc cac ctg ctt tca gaa atg gag ctg atg tgc gac agg atc      1109
Val Ser Ser His Leu Leu Ser Glu Met Glu Leu Met Cys Asp Arg Ile
        190             195             200

gcc att att caa aac gga aag ctc cgg gat att cag cat gtg cac gga      1157
Ala Ile Ile Gln Asn Gly Lys Leu Arg Asp Ile Gln His Val His Gly
205             210             215

ccc gct cgg gat gag aag aag cgc tat tat att caa gcg gac gac acc      1205
Pro Ala Arg Asp Glu Lys Lys Arg Tyr Tyr Ile Gln Ala Asp Asp Thr
220             225             230             235

cag gct ctc aca cgg gaa gcg gct gct ttc aga aag gtg aag gtt gac      1253
Gln Ala Leu Thr Arg Glu Ala Ala Ala Phe Arg Lys Val Lys Val Asp
                240             245             250

gaa gcg gaa ggc ggg ata gag ctc agc att caa aag gat gaa gtg cct      1301
```

```
Glu Ala Glu Gly Gly Ile Glu Leu Ser Ile Gln Lys Asp Glu Val Pro
            255                 260                 265

gat ttg att aaa cac ttg aca gac agc ggt gtt cgc tta tat gaa gtg      1349
Asp Leu Ile Lys His Leu Thr Asp Ser Gly Val Arg Leu Tyr Glu Val
            270                 275                 280

aag gct gtg aac aaa tcg ctg gaa gac cga ttc ctg gaa atc acc gca      1397
Lys Ala Val Asn Lys Ser Leu Glu Asp Arg Phe Leu Glu Ile Thr Ala
            285                 290                 295

gat aag gag gaa gct cag cat gtt taatctcatc gtaaatgaat ggatcaaaat    1451
Asp Lys Glu Glu Ala Gln His Val
300                 305

ttttaatcga aaaggcacat acgtcatgat cggaattttg ctgttagctg tcatcgggct    1511

gggcgttctc acaaagacga tcggagagac agaccaaaac acggactgga aaaaggaatt    1571

ggcgcaggaa ataaggacaa gggggcttag t                                  1602
```

<210> 2

<211> 307

<212> PRT

<213> Bacillus licheniformis


<400> 2

```
Leu Glu Thr Leu Leu Glu Leu Lys Asn Val Ser Lys Thr Ile Arg Gly
1               5                   10                  15

Lys Lys Ile Ile Glu Gly Leu Ser Phe Asp Val Arg Ala Gly Glu Ile
            20                  25                  30

Phe Gly Phe Leu Gly Pro Asn Gly Ala Gly Lys Thr Thr Thr Ile Arg
        35                  40                  45

Met Ile Val Gly His Met Ser Ile Thr Ala Gly Glu Ile Ala Val Cys
    50                  55                  60

Gly Val Ser Val Lys Glu Asn Phe Glu Lys Ala Ala Arg His Ile Gly
65                  70                  75                  80

Ala Ile Val Glu Asn Pro Glu Leu Tyr Lys Phe Leu Thr Gly Tyr Gln
                85                  90                  95

Asn Leu Gln Gln Tyr Ala Arg Met Thr Lys Gly Val Thr Lys Lys Lys
            100                 105                 110

Ile Asp Glu Ile Val Glu Leu Val Gly Leu Lys Asn Arg Ile Asn Asp
            115                 120                 125

Lys Val Lys Thr Tyr Ser Leu Gly Met Arg Gln Arg Leu Gly Leu Ala
            130                 135                 140
```

```
Gln Ser Leu Leu His Asp Pro Lys Leu Leu Ile Leu Asp Glu Pro Thr
145             150             155             160

Asn Gly Leu Asp Pro Ala Gly Ile Arg Glu Ile Arg Asp Tyr Leu Arg
                165             170             175

Lys Leu Thr Arg Glu Lys Gly Met Ala Val Ile Val Ser Ser His Leu
            180             185             190

Leu Ser Glu Met Glu Leu Met Cys Asp Arg Ile Ala Ile Ile Gln Asn
        195             200             205

Gly Lys Leu Arg Asp Ile Gln His Val His Gly Pro Ala Arg Asp Glu
    210             215             220

Lys Lys Arg Tyr Tyr Ile Gln Ala Asp Asp Thr Gln Ala Leu Thr Arg
225             230             235             240

Glu Ala Ala Ala Phe Arg Lys Val Lys Val Asp Glu Ala Glu Gly Gly
            245             250             255

Ile Glu Leu Ser Ile Gln Lys Asp Glu Val Pro Asp Leu Ile Lys His
            260             265             270

Leu Thr Asp Ser Gly Val Arg Leu Tyr Glu Val Lys Ala Val Asn Lys
        275             280             285

Ser Leu Glu Asp Arg Phe Leu Glu Ile Thr Ala Asp Lys Glu Glu Ala
    290             295             300

Gln His Val
305


<210>   3

<211>   1938

<212>   DNA

<213>   Bacillus licheniformis


<220>

<221>   CDS

<222>   (501)..(1457)

<223>


<400>   3
cgcatcggac ttttgaacgt cggaaccgaa gataaaaaag gaaatgagct tgccaagcag      60
```

Page 4

```
acctttcaaa aattgaagga aaccgatttg aatttcatcg gcaatgtgga agcccgcgat     120

atgctggacg gagtcgctga tgtcatcgtc acagacggct ttaccggtaa cgttgccttg     180

aaaacggtcg agggcgcggc gctgtccatt tttaaaatgc tgagaacgac gctgacttcg     240

agcttcacgg cgaagctcgc cgcttctgca ctgaagccga agctgaaaga aatgaaaacg     300

aaaatggatt actctgaata cggcggagcc ggattgttcg gcttaaaggc gcccgtcatc     360

aaagcgcacg gatcatctga cggacgcgcc gtttatcacg cgatccgcca ggccagagag     420

atggtcagcc aaaatgtcgc ggcatttatc gaagaaaaaa ttcaacaaaa agcagatgaa     480

tagtctggag gttttaacac atg ggc aag att gct ttt cta ttc ccg ggc caa     533
                         Met Gly Lys Ile Ala Phe Leu Phe Pro Gly Gln
                          1               5                   10

ggt tcg cag cat atc ggc atg gga cac gaa ttg tat gaa aaa gaa ccg      581
Gly Ser Gln His Ile Gly Met Gly His Glu Leu Tyr Glu Lys Glu Pro
             15              20                  25

aat gcg aag aag att ttt gaa gaa gcg gat caa acg ctt gaa aca aaa      629
Asn Ala Lys Lys Ile Phe Glu Glu Ala Asp Gln Thr Leu Glu Thr Lys
         30                  35                  40

ctg agc acc ctc atg ttt gaa ggg gat gca aag gaa ctg acg ctt aca      677
Leu Ser Thr Leu Met Phe Glu Gly Asp Ala Lys Glu Leu Thr Leu Thr
     45                  50                  55

tac aac gcg cag cca agc ctt tta acg gcg agc atc gca gcg ctt gaa      725
Tyr Asn Ala Gln Pro Ser Leu Leu Thr Ala Ser Ile Ala Ala Leu Glu
60                  65                  70                  75

aaa ctg aag gaa tac ggc att aaa gcc gac tat gcg gca ggt cac agc      773
Lys Leu Lys Glu Tyr Gly Ile Lys Ala Asp Tyr Ala Ala Gly His Ser
                 80                  85                  90

ctc ggc gaa tac agc gca ttg gtc gct gcc ggc gcc ttg tcg ttt aaa     821
Leu Gly Glu Tyr Ser Ala Leu Val Ala Ala Gly Ala Leu Ser Phe Lys
             95                  100                 105

gat gcg gtt tat gcc gtc aga aag cgc ggc gaa ttc atg aat gaa gcc     869
Asp Ala Val Tyr Ala Val Arg Lys Arg Gly Glu Phe Met Asn Glu Ala
         110                 115                 120

gtg ccg gcg gga gaa ggc gcg atg gcg gcc att ctc ggc atg gac agc     917
Val Pro Ala Gly Glu Gly Ala Met Ala Ala Ile Leu Gly Met Asp Ser
     125                 130                 135

cag gcg ctg aaa gaa gtg acg gac aaa att tcc gaa gaa gga aac ctt     965
Gln Ala Leu Lys Glu Val Thr Asp Lys Ile Ser Glu Glu Gly Asn Leu
140                 145                 150                 155

gtt cag ctc gcc aat ttg aac tgc cct ggg caa atc gtc atc tcg gga    1013
Val Gln Leu Ala Asn Leu Asn Cys Pro Gly Gln Ile Val Ile Ser Gly
                 160                 165                 170

aca gct aaa ggc gtg gag ctc gct tca gag ctt gcg aaa gaa aag ggc    1061
Thr Ala Lys Gly Val Glu Leu Ala Ser Glu Leu Ala Lys Glu Lys Gly
             175                 180                 185

gca aaa cgc gcg att cct ctc gaa gtc agc ggg ccg ttc cat tct gag    1109
Ala Lys Arg Ala Ile Pro Leu Glu Val Ser Gly Pro Phe His Ser Glu
         190                 195                 200

ctg atg aag ccg gca gct gat aag ctt cgt gaa gtt ctt gat gcg tgc    1157
Leu Met Lys Pro Ala Ala Asp Lys Leu Arg Glu Val Leu Asp Ala Cys
```

```
      Leu Met Lys Pro Ala Ala Asp Lys Leu Arg Glu Val Leu Asp Ala Cys
          205             210             215

      acg atc aac gac gca gcc att ccg gtc gtc tcc aac gta acg gcc gac      1205
      Thr Ile Asn Asp Ala Ala Ile Pro Val Val Ser Asn Val Thr Ala Asp
      220             225             230             235

      ttt gta acg gat aaa gac gac att aag aat aaa ctg att gaa cag ctg      1253
      Phe Val Thr Asp Lys Asp Asp Ile Lys Asn Lys Leu Ile Glu Gln Leu
              240             245             250

      tat tcc cct gta cgc ttt gaa gaa aca atc agc cgc ctg att gac gaa      1301
      Tyr Ser Pro Val Arg Phe Glu Glu Thr Ile Ser Arg Leu Ile Asp Glu
              255             260             265

      ggc gtc acg acc ttc att gaa atc ggt ccc gga aag gtt ttg tca ggg      1349
      Gly Val Thr Thr Phe Ile Glu Ile Gly Pro Gly Lys Val Leu Ser Gly
              270             275             280

      ctt gtg aag aaa gtg aac cgc aga gtc aaa acg att gct gta tca gac      1397
      Leu Val Lys Lys Val Asn Arg Arg Val Lys Thr Ile Ala Val Ser Asp
          285             290             295

      ccg aac aca att gaa ctt gcc gtt caa acg ttg aag gag gaa aac gaa      1445
      Pro Asn Thr Ile Glu Leu Ala Val Gln Thr Leu Lys Glu Glu Asn Glu
      300             305             310             315

      aat gct gga aaa taaaacagcc gttgtgacag gagcctcaag aggaatcggc         1497
      Asn Ala Gly Lys
```

cgcgcgatcg ccctggacct ggcgaaaaac ggagcaaatg tcgtcgtcaa ctacgcggga      1557

aatgaagcga aagcgaacga agtcgtagac gaaatcaaag cgctcggccg cgatgcgttt      1617

gcttttaaag cggacgtttc caatgcggat gaggttcagg cgatgatgaa ggaagcggtc      1677

ggacgcttcg gcacgcttga catccttgtc aacaatgcgg gcattactaa agacaatctg      1737

ttcatgagaa tgaaagaaga tgaatgggac gacgtcatta acataaactt aaaaggtgtg      1797

ttcaattgtt caaaagctgt gacaagacag atgatgaaac aaagaagcgg ccggatcatc      1857

aatatcacct cggttgtagg cgtcgtcggt aacgccgggc aggccaacta tgtcgcggct      1917

aaatcaggcg tattccagta g                                               1938

<210> 4

<211> 319

<212> PRT

<213> Bacillus licheniformis


<400> 4

```
Met Gly Lys Ile Ala Phe Leu Phe Pro Gly Gln Gly Ser Gln His Ile
1               5               10              15


Gly Met Gly His Glu Leu Tyr Glu Lys Glu Pro Asn Ala Lys Lys Ile
        20              25              30
```

Phe Glu Glu Ala Asp Gln Thr Leu Glu Thr Lys Leu Ser Thr Leu Met
        35              40              45

Phe Glu Gly Asp Ala Lys Glu Leu Thr Leu Thr Tyr Asn Ala Gln Pro
        50              55              60

Ser Leu Leu Thr Ala Ser Ile Ala Ala Leu Glu Lys Leu Lys Glu Tyr
65              70              75              80

Gly Ile Lys Ala Asp Tyr Ala Ala Gly His Ser Leu Gly Glu Tyr Ser
                85              90              95

Ala Leu Val Ala Ala Gly Ala Leu Ser Phe Lys Asp Ala Val Tyr Ala
            100             105             110

Val Arg Lys Arg Gly Glu Phe Met Asn Glu Ala Val Pro Ala Gly Glu
        115             120             125

Gly Ala Met Ala Ala Ile Leu Gly Met Asp Ser Gln Ala Leu Lys Glu
130             135             140

Val Thr Asp Lys Ile Ser Glu Glu Gly Asn Leu Val Gln Leu Ala Asn
145             150             155             160

Leu Asn Cys Pro Gly Gln Ile Val Ile Ser Gly Thr Ala Lys Gly Val
                165             170             175

Glu Leu Ala Ser Glu Leu Ala Lys Glu Lys Gly Ala Lys Arg Ala Ile
            180             185             190

Pro Leu Glu Val Ser Gly Pro Phe His Ser Glu Leu Met Lys Pro Ala
        195             200             205

Ala Asp Lys Leu Arg Glu Val Leu Asp Ala Cys Thr Ile Asn Asp Ala
210             215             220

Ala Ile Pro Val Val Ser Asn Val Thr Ala Asp Phe Val Thr Asp Lys
225             230             235             240

Asp Asp Ile Lys Asn Lys Leu Ile Glu Gln Leu Tyr Ser Pro Val Arg
                245             250             255

Phe Glu Glu Thr Ile Ser Arg Leu Ile Asp Glu Gly Val Thr Thr Phe
            260             265             270

Ile Glu Ile Gly Pro Gly Lys Val Leu Ser Gly Leu Val Lys Lys Val
        275             280             285

Asn Arg Arg Val Lys Thr Ile Ala Val Ser Asp Pro Asn Thr Ile Glu
290             295             300

Leu Ala Val Gln Thr Leu Lys Glu Glu·Asn Glu Asn Ala Gly Lys
305                        310                    315

<210>  5

<211>  2119

<212>  DNA

<213>  Bacillus licheniformis


<220>

<221>  CDS

<222>  (381)..(1616)

<223>


<400>  5
tagtggatca tacaatcatg aatggcagag aagtgttcaa gtttgccgtc cgccaaatgg     60

gtgaatccag tgtaaatgtt attgaaaaag cgggattatc caaagaggat gtcgactttt    120

tggtgccgca ccaagccaac atccggatca tggaagcggc gcgcgaacgt ttggagcttc    180

cggtcgaaaa aatgtcgaaa acggttcata atacggaaa cacatcagct gcttctattc     240

cgatttctct tgtggaagag cttgaagctg gtaaaataaa agacggtgac gtcattgtga    300

tggtcggctt cggcggaggt ttgacatggg gagccatcgc catgcgctgg ggccgctaaa    360

agaaaggtga ggtgcaactt atg aat agt aaa aga gta gtc att acc gga tta     413
                        Met Asn Ser Lys Arg Val Val Ile Thr Gly Leu
                         1               5                  10

gga gct tta tca cct ctt ggt aat gac gtc aaa aca agc tgg aac aac       461
Gly Ala Leu Ser Pro Leu Gly Asn Asp Val Lys Thr Ser Trp Asn Asn
                15                  20                  25

gcg atc aac ggg gtt tcc gga gtc ggc ccg ata aca agg gtg aac gcg       509
Ala Ile Asn Gly Val Ser Gly Val Gly Pro Ile Thr Arg Val Asn Ala
            30                  35                  40

gac gag tat cct gct aaa gta gcc gca gaa ctg aaa gat ttt aag ata       557
Asp Glu Tyr Pro Ala Lys Val Ala Ala Glu Leu Lys Asp Phe Lys Ile
        45                  50                  55

gaa gat tat atg gat aaa aaa gaa gcg aga aaa atg gac cgc ttc aca       605
Glu Asp Tyr Met Asp Lys Lys Glu Ala Arg Lys Met Asp Arg Phe Thr
60                  65                  70                  75

caa tat gcc gtt gtg tcg gcg aaa atg gcc gtg gaa gac gcg ggc ctc       653
Gln Tyr Ala Val Val Ser Ala Lys Met Ala Val Glu Asp Ala Gly Leu
                80                  85                  90

gac att aac gaa agc aac gca gac cgg atc ggc gtc tgg gtc ggt tcc       701
Asp Ile Asn Glu Ser Asn Ala Asp Arg Ile Gly Val Trp Val Gly Ser
                95                  100                 105

gga atc ggc ggc ctt gaa acg ctt gaa cag cag ttc gaa aca ttt tta       749
Gly Ile Gly Gly Leu Glu Thr Leu Glu Gln Gln Phe Glu Thr Phe Leu
            110                 115                 120

```
aca aaa gga ccg cgg cgc gtc agc ccg ttc ttt gta ccg atg atg att        797
Thr Lys Gly Pro Arg Arg Val Ser Pro Phe Phe Val Pro Met Met Ile
    125             130                 135

cct gat atg gcg acc ggc cag att tcg atc tct ctc ggt gca aaa ggg        845
Pro Asp Met Ala Thr Gly Gln Ile Ser Ile Ser Leu Gly Ala Lys Gly
140             145                 150                 155

gtc aac tct tgt acg gtg acg gct tgc gcg acg ggt aca aac tca atc       893
Val Asn Ser Cys Thr Val Thr Ala Cys Ala Thr Gly Thr Asn Ser Ile
                160                 165                 170

ggc gac gct ttc aag gtc att cag cgc ggc gat gcc gat gtc atg atc       941
Gly Asp Ala Phe Lys Val Ile Gln Arg Gly Asp Ala Asp Val Met Ile
            175                 180                 185

tcc gga gga acg gaa gcc ccg ctt aca aga atg tca ttt gcc ggc ttt       989
Ser Gly Gly Thr Glu Ala Pro Leu Thr Arg Met Ser Phe Ala Gly Phe
        190                 195                 200

tca gcc aat aaa gcg cta tcc act aat ccg gat ccg aaa acg gca agc      1037
Ser Ala Asn Lys Ala Leu Ser Thr Asn Pro Asp Pro Lys Thr Ala Ser
    205                 210                 215

cgt ccg ttt gac aaa aac cgc gac ggc ttc gtc atg gga gaa gga gct      1085
Arg Pro Phe Asp Lys Asn Arg Asp Gly Phe Val Met Gly Glu Gly Ala
220                 225                 230                 235

gga atc gtc gtt ctt gag gag ctt gag cac gcc tta aac cgc gga gcg      1133
Gly Ile Val Val Leu Glu Glu Leu Glu His Ala Leu Asn Arg Gly Ala
            240                 245                 250

aca atc tac gcg gaa gtg gtc ggc tac gga tca acg gga gac gcc tat      1181
Thr Ile Tyr Ala Glu Val Val Gly Tyr Gly Ser Thr Gly Asp Ala Tyr
        255                 260                 265

cat att acg gca cct gcg cca aat gga gaa ggc ggc gtc aga gcc atg      1229
His Ile Thr Ala Pro Ala Pro Asn Gly Glu Gly Gly Val Arg Ala Met
    270                 275                 280

aaa gaa gcg atc aaa gat gcg ggt ctg tcc cct gtg gac att gat tat      1277
Lys Glu Ala Ile Lys Asp Ala Gly Leu Ser Pro Val Asp Ile Asp Tyr
    285                 290                 295

atc aat gct cat gga aca agc acg ccg tac aat gat aaa ttt gaa acg      1325
Ile Asn Ala His Gly Thr Ser Thr Pro Tyr Asn Asp Lys Phe Glu Thr
300             305                 310                 315

atc gcc att aaa gaa gtg ttc gga gag cat gcc tac aag ctt gcc gtc      1373
Ile Ala Ile Lys Glu Val Phe Gly Glu His Ala Tyr Lys Leu Ala Val
            320                 325                 330

agc tcg aca aaa tcg atg aca ggg cac ttg ctc ggg gct gcc gga ggc      1421
Ser Ser Thr Lys Ser Met Thr Gly His Leu Leu Gly Ala Ala Gly Gly
            335                 340                 345

atc gaa gcg att ttc tcc gtc ctt gcc att aag gaa ggc atc att cct      1469
Ile Glu Ala Ile Phe Ser Val Leu Ala Ile Lys Glu Gly Ile Ile Pro
        350                 355                 360

ccg aca atc aat atc gaa acg cct gat gaa gat tgc gat ctg gac tat      1517
Pro Thr Ile Asn Ile Glu Thr Pro Asp Glu Asp Cys Asp Leu Asp Tyr
    365                 370                 375

gta ccg gat caa gcc cgc cgc cag gac gtg aat gtt gtg ctc agc aac      1565
Val Pro Asp Gln Ala Arg Arg Gln Asp Val Asn Val Val Leu Ser Asn
380                 385                 390                 395
```

Page 9

```
tcg ctt ggt ttc ggc gga cac aac gca acg ctg atc ttc aaa aaa tat      1613
Ser Leu Gly Phe Gly Gly His Asn Ala Thr Leu Ile Phe Lys Lys Tyr
            400                     405                     410

gaa taacccgcaa tataaacggc tgccatttta aatggtggcc gttttttta            1666
Glu


gcgtgcatac catacagatg aaaaaggagt gtgatctaaa atgagtatag caagcacgta    1726

tgaatggctt gaacaggctt catccgtgcg ggaattgcca agctttctcc tgccgtattt    1786

taaaggcgct tcaaaagcgg atgcggcggc gattgtcagc cacctgcaga tccacggctt    1846

gttccgttcg tggacggacg gcgaacgtac gatgaagcag cttaaggaaa acggcgtatt    1906

tcagcatgtc agacgggaag aacagctgct ctgcaaaaga tggaacgggc cggatgttcc    1966

gattttcatc ctgccggtcg atgaaagaaa ccgcaaaatc cgcgtcgaat cggctcgaa     2026

atcaggcctc gcgttcagcg acaaattatt tttgttttta tcggccgatc ttccccggtc    2086

ttcaatttct gcgatcatga cacacgaata taa                                 2119
```

<210> 6

<211> 412

<212> PRT

<213> Bacillus licheniformis


<400> 6

```
Met Asn Ser Lys Arg Val Val Ile Thr Gly Leu Gly Ala Leu Ser Pro
1               5                   10                  15

Leu Gly Asn Asp Val Lys Thr Ser Trp Asn Asn Ala Ile Asn Gly Val
                20                  25                  30

Ser Gly Val Gly Pro Ile Thr Arg Val Asn Ala Asp Glu Tyr Pro Ala
            35                  40                  45

Lys Val Ala Ala Glu Leu Lys Asp Phe Lys Ile Glu Asp Tyr Met Asp
        50                  55                  60

Lys Lys Glu Ala Arg Lys Met Asp Arg Phe Thr Gln Tyr Ala Val Val
65                  70                  75                  80

Ser Ala Lys Met Ala Val Glu Asp Ala Gly Leu Asp Ile Asn Glu Ser
                85                  90                  95

Asn Ala Asp Arg Ile Gly Val Trp Val Gly Ser Gly Ile Gly Gly Leu
                100                 105                 110

Glu Thr Leu Glu Gln Gln Phe Glu Thr Phe Leu Thr Lys Gly Pro Arg
            115                 120                 125
```

```
Arg Val Ser Pro Phe Phe Val Pro Met Met Ile Pro Asp Met Ala Thr
    130             135             140
```

```
Gly Gln Ile Ser Ile Ser Leu Gly Ala Lys Gly Val Asn Ser Cys Thr
145             150             155             160
```

```
Val Thr Ala Cys Ala Thr Gly Thr Asn Ser Ile Gly Asp Ala Phe Lys
            165             170             175
```

```
Val Ile Gln Arg Gly Asp Ala Asp Val Met Ile Ser Gly Gly Thr Glu
            180             185             190
```

```
Ala Pro Leu Thr Arg Met Ser Phe Ala Gly Phe Ser Ala Asn Lys Ala
        195             200             205
```

```
Leu Ser Thr Asn Pro Asp Pro Lys Thr Ala Ser Arg Pro Phe Asp Lys
    210             215             220
```

```
Asn Arg Asp Gly Phe Val Met Gly Glu Gly Ala Gly Ile Val Val Leu
225             230             235             240
```

```
Glu Glu Leu Glu His Ala Leu Asn Arg Gly Ala Thr Ile Tyr Ala Glu
            245             250             255
```

```
Val Val Gly Tyr Gly Ser Thr Gly Asp Ala Tyr His Ile Thr Ala Pro
            260             265             270
```

```
Ala Pro Asn Gly Glu Gly Gly Val Arg Ala Met Lys Glu Ala Ile Lys
        275             280             285
```

```
Asp Ala Gly Leu Ser Pro Val Asp Ile Asp Tyr Ile Asn Ala His Gly
    290             295             300
```

```
Thr Ser Thr Pro Tyr Asn Asp Lys Phe Glu Thr Ile Ala Ile Lys Glu
305             310             315             320
```

```
Val Phe Gly Glu His Ala Tyr Lys Leu Ala Val Ser Ser Thr Lys Ser
            325             330             335
```

```
Met Thr Gly His Leu Leu Gly Ala Ala Gly Gly Ile Glu Ala Ile Phe
            340             345             350
```

```
Ser Val Leu Ala Ile Lys Glu Gly Ile Ile Pro Pro Thr Ile Asn Ile
        355             360             365
```

```
Glu Thr Pro Asp Glu Asp Cys Asp Leu Asp Tyr Val Pro Asp Gln Ala
    370             375             380
```

```
Arg Arg Gln Asp Val Asn Val Val Leu Ser Asn Ser Leu Gly Phe Gly
385             390             395             400
```

```
Gly His Asn Ala Thr Leu Ile Phe Lys Lys Tyr Glu
                405                 410
```

<210> 7

<211> 8158

<212> DNA

<213> Bacillus licheniformis


<220>

<221> CDS

<222> (501)..(7655)

<223>


<400> 7

```
atcaccggaa tctatgccca tatcggctgc ctgatgacgg catgtgatgc ttttatgcaa      60

gatattgaaa catttttcgt cggagatgcg gtggcggatt tctcattaaa gcatcatgaa     120

atggccatca catacgcagc tgaccgttgc agtgtcacaa ttactgcgga acagcttctg     180

agccgactga aagatggcga aagccatgaa cacgcaaaag aaagctctca tgcactgaca     240

aaacagatta ttcgcgaaca ggtggcatcg cttcttcttg aatcgccgga aaagataagc     300

gacagcgaaa atttgatcta ccggggaatt gattccgtga gaatcatgag tctcgcagag     360

cgctggcgcc gagcggggcg ggaggtgtcg tttgtagagc ttgcggagga cccttcgatt     420

gaaaactggt ggagactgtt gtcctcctct aaagaagcac ctttgccaaa tgcagactat     480

caatgaagga ggtcacctaa atg cct gaa tgt caa cat aac cga aag cca tta     533
                        Met Pro Glu Cys Gln His Asn Arg Lys Pro Leu
                        1                5                    10

tca gga gcg caa gcc ggg att tgg ttt gct cag cag ctt gat ccg gaa     581
Ser Gly Ala Gln Ala Gly Ile Trp Phe Ala Gln Gln Leu Asp Pro Glu
            15              20                  25

aat ccg atc tac aat aca gct gaa tac gtt gaa att aaa ggc ccg ctt     629
Asn Pro Ile Tyr Asn Thr Ala Glu Tyr Val Glu Ile Lys Gly Pro Leu
        30              35                  40

gat cag gag ctt ttc gaa aaa gcg ctg cgc cat gtc att aaa gaa gct     677
Asp Gln Glu Leu Phe Glu Lys Ala Leu Arg His Val Ile Lys Glu Ala
    45              50                  55

gaa tca ttt cat gcc aga ttt gga gaa gat caa gac gga cca tgg caa     725
Glu Ser Phe His Ala Arg Phe Gly Glu Asp Gln Asp Gly Pro Trp Gln
60              65                  70                  75

gag atc gtt ccg tca aca gat ttt ccg cta cat tac att gat gtc agc     773
Glu Ile Val Pro Ser Thr Asp Phe Pro Leu His Tyr Ile Asp Val Ser
            80                  85                  90

tca gaa acc gat ccg gaa cag gcg gct aaa agc tgg atg atg gat gac     821
Ser Glu Thr Asp Pro Glu Gln Ala Ala Lys Ser Trp Met Met Asp Asp
```

```
Ser Glu Thr Asp Pro Glu Gln Ala Ala Lys Ser Trp Met Met Asp Asp
        95              100             105

ctt gcc cgt ccg gtt gat ctg act cgc ggc ccg ctt ttt aca gaa gcg    869
Leu Ala Arg Pro Val Asp Leu Thr Arg Gly Pro Leu Phe Thr Glu Ala
        110             115             120

ctt ttt aaa gcg gcg caa gat cat tac ttc tgg tat cag cga act cac    917
Leu Phe Lys Ala Ala Gln Asp His Tyr Phe Trp Tyr Gln Arg Thr His
        125             130             135

cat atc gca aca gac ggg ttc agc ttt aca ttg atc gcc gag cgg ctg    965
His Ile Ala Thr Asp Gly Phe Ser Phe Thr Leu Ile Ala Glu Arg Leu
140             145             150             155

tca aaa ata tat acc gca ttg atg cag aac aaa tcg atc gac cag agc   1013
Ser Lys Ile Tyr Thr Ala Leu Met Gln Asn Lys Ser Ile Asp Gln Ser
                160             165             170

gga gcc ttt ggc tcc ctc gat ttg att ctc gcg gag gaa acg gct tac   1061
Gly Ala Phe Gly Ser Leu Asp Leu Ile Leu Ala Glu Glu Thr Ala Tyr
                175             180             185

cgt gca tca gaa cag tat cag gaa gat cgg cga ttt tgg ctt ggg cgt   1109
Arg Ala Ser Glu Gln Tyr Gln Glu Asp Arg Arg Phe Trp Leu Gly Arg
                190             195             200

ttc agc gat gaa ccg gag gtg gtc agc ctt gcc gaa agg gcg ccg cgt   1157
Phe Ser Asp Glu Pro Glu Val Val Ser Leu Ala Glu Arg Ala Pro Arg
        205             210             215

acc tct tca agt ttt ctt cgc agg tct gaa cac ctg cca agt gag gat   1205
Thr Ser Ser Ser Phe Leu Arg Arg Ser Glu His Leu Pro Ser Glu Asp
220             225             230             235

gct gat cgt ctt ctg tct gcc gcg agc aga atg ggg gca act tgg cac   1253
Ala Asp Arg Leu Leu Ser Ala Ala Ser Arg Met Gly Ala Thr Trp His
                240             245             250

gaa acc gtg atg gca gca gca gca ata tat gtt cac cgt ctg aca ggg   1301
Glu Thr Val Met Ala Ala Ala Ala Ile Tyr Val His Arg Leu Thr Gly
                255             260             265

gca aat gac gtc gtt ctc ggg atg ccg atg atg tgc cgc ctc ggg tca   1349
Ala Asn Asp Val Val Leu Gly Met Pro Met Met Cys Arg Leu Gly Ser
                270             275             280

gcc gct ctt cat att ccg gga atg gtc atg aat ctc ctc ccg ctt cga   1397
Ala Ala Leu His Ile Pro Gly Met Val Met Asn Leu Leu Pro Leu Arg
        285             290             295

att ggc atg cag ccg caa atg agc ata gga gag cta gtc aag caa atc   1445
Ile Gly Met Gln Pro Gln Met Ser Ile Gly Glu Leu Val Lys Gln Ile
300             305             310             315

tcg ggt gag atg atg aag ctg cgg cgc cat cag cat tac cgc cac gaa   1493
Ser Gly Glu Met Met Lys Leu Arg Arg His Gln His Tyr Arg His Glu
                320             325             330

gaa ttg cgc cgg gat ctc aag ctg ctt ggc gaa aac cag cgg ctg ttc   1541
Glu Leu Arg Arg Asp Leu Lys Leu Leu Gly Glu Asn Gln Arg Leu Phe
                335             340             345

ggt ccg cag ctt aat ttg atg cct ttt gag aac cgc ttg aat ttt gcc   1589
Gly Pro Gln Leu Asn Leu Met Pro Phe Glu Asn Arg Leu Asn Phe Ala
                350             355             360

ggc tgt caa ggc atc gtg cat aat ctt gcc acg gga cct gtg gac gat   1637
```

Page 13

```
Gly Cys Gln Gly Ile Val His Asn Leu Ala Thr Gly Pro Val Asp Asp
365                 370             375

ttg tca atc aac att tac ggc cgt ccg gac ggt ggc ggg cta aaa gta        1685
Leu Ser Ile Asn Ile Tyr Gly Arg Pro Asp Gly Gly Gly Leu Lys Val
380                 385             390             395

aac atg gat gca aac cct gct gtt tat cat gcg gat gaa ctt gaa gat        1733
Asn Met Asp Ala Asn Pro Ala Val Tyr His Ala Asp Glu Leu Glu Asp
                400             405             410

cac ggc aat cgt ttt ctt act ctt ctg aaa acg att gcc gtt tgc gaa        1781
His Gly Asn Arg Phe Leu Thr Leu Leu Lys Thr Ile Ala Val Cys Glu
            415             420             425

caa act cag ccg gtg ggg acg ttg gac att ctt ctc cct gaa gaa cgc        1829
Gln Thr Gln Pro Val Gly Thr Leu Asp Ile Leu Leu Pro Glu Glu Arg
            430             435             440

gct caa ata ttg att gaa tgg aat caa aca gaa cac ggg ctt ccc aaa        1877
Ala Gln Ile Leu Ile Glu Trp Asn Gln Thr Glu His Gly Leu Pro Lys
        445             450             455

gag agt ctt ccg cag cga ttt gaa agg atg gca aag gag tgt cct gaa        1925
Glu Ser Leu Pro Gln Arg Phe Glu Arg Met Ala Lys Glu Cys Pro Glu
460             465             470             475

tct ccg gca gtc gtt tgc aat gat aag gtt ctc act tat tcc gaa tta        1973
Ser Pro Ala Val Val Cys Asn Asp Lys Val Leu Thr Tyr Ser Glu Leu
                480             485             490

aat caa aag gct aat caa ctg gct cac ttg ctg atc gat cag ggg gcg        2021
Asn Gln Lys Ala Asn Gln Leu Ala His Leu Leu Ile Asp Gln Gly Ala
            495             500             505

aaa ccc gag aca ttt att gcc ttg gcg ttg ccg cgg tca gcg gaa atg        2069
Lys Pro Glu Thr Phe Ile Ala Leu Ala Leu Pro Arg Ser Ala Glu Met
            510             515             520

gtt gtc agc atg ctc gcc gtt ctc aaa gct gga gcg gct tat tta cct        2117
Val Val Ser Met Leu Ala Val Leu Lys Ala Gly Ala Ala Tyr Leu Pro
        525             530             535

ata gat ccc gat tat ccg gct gac agg atc gag tat atg ctc aac gat        2165
Ile Asp Pro Asp Tyr Pro Ala Asp Arg Ile Glu Tyr Met Leu Asn Asp
540             545             550             555

gca cag ccg ctg ctt gtt atg acc agc aag gaa gcg caa gac acg atc        2213
Ala Gln Pro Leu Leu Val Met Thr Ser Lys Glu Ala Gln Asp Thr Ile
                560             565             570

ggt tcg caa atg ccg aag tta atc ctc gat gaa caa act gtt atg gag        2261
Gly Ser Gln Met Pro Lys Leu Ile Leu Asp Glu Gln Thr Val Met Glu
            575             580             585

aaa atg tcc ggc tgt tct gaa gaa aat ccc ggc gaa cag cat tcc ggc        2309
Lys Met Ser Gly Cys Ser Glu Glu Asn Pro Gly Glu Gln His Ser Gly
            590             595             600

aac cag ccg gca tat atg att tat acg tcg gga tca acc ggc aga cct        2357
Asn Gln Pro Ala Tyr Met Ile Tyr Thr Ser Gly Ser Thr Gly Arg Pro
            605             610             615

aaa ggc gtg gtc gta caa gct gaa agc tta ttc aat ttt ttg ctg tca        2405
Lys Gly Val Val Val Gln Ala Glu Ser Leu Phe Asn Phe Leu Leu Ser
620             625             630             635

atg cag gac atg ttc gcg ctt aat caa gat gac cgg ctg ctg gcc gtc        2453
```

```
Met Gln Asp Met Phe Ala Leu Asn Gln Asp Asp Arg Leu Leu Ala Val
            640             645             650

act aca att gca ttc gac atc tca gca ctc gaa att tat ttg ccg ctg     2501
Thr Thr Ile Ala Phe Asp Ile Ser Ala Leu Glu Ile Tyr Leu Pro Leu
            655             660             665

atc agc gga tca gca gtc gtt ctt gcc gag aag gaa acg gta caa gat     2549
Ile Ser Gly Ser Ala Val Val Leu Ala Glu Lys Glu Thr Val Gln Asp
            670             675             680

ccg tcc gaa ttg gcc aaa atg att gaa aca tac gaa att aca ata atg     2597
Pro Ser Glu Leu Ala Lys Met Ile Glu Thr Tyr Glu Ile Thr Ile Met
            685             690             695

cag gct aca ccg acc ctc tgg cat gca ttg gcc tcg agc gcc ccg gaa     2645
Gln Ala Thr Pro Thr Leu Trp His Ala Leu Ala Ser Ser Ala Pro Glu
700             705             710             715

aag ctc aaa ggt ctt cgt gcg ctt gtc ggc ggc gaa gct ttg caa tcc     2693
Lys Leu Lys Gly Leu Arg Ala Leu Val Gly Gly Glu Ala Leu Gln Ser
            720             725             730

agc ctg gcc cgg caa ttg cag cag ctc gcg tgt tct tta acg aac ctt     2741
Ser Leu Ala Arg Gln Leu Gln Gln Leu Ala Cys Ser Leu Thr Asn Leu
            735             740             745

tat gga ccg acg gaa aca aca att tgg tct aca gcc gct gcg ctg gaa     2789
Tyr Gly Pro Thr Glu Thr Thr Ile Trp Ser Thr Ala Ala Ala Leu Glu
            750             755             760

ggc aac tgt acg gaa cct ccg gga atc ggc tgt gcg att tgg aat acg     2837
Gly Asn Cys Thr Glu Pro Pro Gly Ile Gly Cys Ala Ile Trp Asn Thr
            765             770             775

cag ctt tat gtc ctg gac gcc gga tta cag cca gtg cct ccg gga acg     2885
Gln Leu Tyr Val Leu Asp Ala Gly Leu Gln Pro Val Pro Pro Gly Thr
780             785             790             795

gct gga gaa ctt tat gtg gca gga aca ggg gta gcg cgc ggc tac ttg     2933
Ala Gly Glu Leu Tyr Val Ala Gly Thr Gly Val Ala Arg Gly Tyr Leu
            800             805             810

aac cga cat gct ctc acg gct gag cgc ttt att gcg aat cca tac ggg     2981
Asn Arg His Ala Leu Thr Ala Glu Arg Phe Ile Ala Asn Pro Tyr Gly
            815             820             825

ccg cct gga tcg cgg atg tac cgg aca ggc gac atc gtg cgc tgg cgc     3029
Pro Pro Gly Ser Arg Met Tyr Arg Thr Gly Asp Ile Val Arg Trp Arg
            830             835             840

gaa gac ggg tcg ctt gat tat atc ggc cgt gca gac cat caa gtc aaa     3077
Glu Asp Gly Ser Leu Asp Tyr Ile Gly Arg Ala Asp His Gln Val Lys
            845             850             855

att cgg ggg ttc cga att gaa atg gga gaa ata gaa gcg gtg ctt gca     3125
Ile Arg Gly Phe Arg Ile Glu Met Gly Glu Ile Glu Ala Val Leu Ala
860             865             870             875

aat cat ccg gtt gtt aaa caa gct gct gct atc gtt cgt gaa gac cag     3173
Asn His Pro Val Val Lys Gln Ala Ala Ala Ile Val Arg Glu Asp Gln
            880             885             890

ccc ggt gac ccg cgt tta ttc gcg tat gcc gtt ccc gct tcg gga gaa     3221
Pro Gly Asp Pro Arg Leu Phe Ala Tyr Ala Val Pro Ala Ser Gly Glu
            895             900             905

agc ctt gat cct gcc gag ctt cgc cgc ttc gtt ggt gag aca ctg ccc     3269
```

```
Ser Leu Asp Pro Ala Glu Leu Arg Arg Phe Val Gly Glu Thr Leu Pro
    910                     915                 920

gac tat atg att ccg tct gca ttt gtc ata ctt gat gaa ctg ccg ctt     3317
Asp Tyr Met Ile Pro Ser Ala Phe Val Ile Leu Asp Glu Leu Pro Leu
    925                     930                 935

acg ccg aac gga aag ctt gac cgc aaa tcc ctg ccg gct ccg gcg gtg     3365
Thr Pro Asn Gly Lys Leu Asp Arg Lys Ser Leu Pro Ala Pro Ala Val
940                     945                 950                 955

agt atg cat aca ggc gga cgc gag ccg agg act ccg caa gag gaa att     3413
Ser Met His Thr Gly Gly Arg Glu Pro Arg Thr Pro Gln Glu Glu Ile
                    960                 965                 970

ttg tgt gat tta ttt gct gag gtg ctg ggg gtg ccg cga gtc agt att     3461
Leu Cys Asp Leu Phe Ala Glu Val Leu Gly Val Pro Arg Val Ser Ile
                975                 980                 985

gat gac agc ttt ttt gac ctc ggc gga cat tcc ctt ctg  gca ggc agg     3509
Asp Asp Ser Phe Phe Asp Leu Gly Gly His Ser Leu Leu  Ala Gly Arg
                990                 995             1000

ctc gtc  ggc cgc att cgg gaa  atg ctc ggc gtc gaa  ctc gga atc       3554
Leu Val  Gly Arg Ile Arg Glu  Met Leu Gly Val Glu  Leu Gly Ile
    1005                1010                1015

ggc cgt  tta ttc gat gag ccc  aca gcc gcc gga ctt  gcc aaa cag       3599
Gly Arg  Leu Phe Asp Glu Pro  Thr Ala Ala Gly Leu  Ala Lys Gln
    1020                1025                1030

ctt gat  cag gcg cag agc gcc  cgt ccg gca ttg cgc  aaa aga gag       3644
Leu Asp  Gln Ala Gln Ser Ala  Arg Pro Ala Leu Arg  Lys Arg Glu
    1035                1040                1045

cgc cgc  aag gag att ccg ctg  tcc ttc gca cag cgg  cgc cta tgg       3689
Arg Arg  Lys Glu Ile Pro Leu  Ser Phe Ala Gln Arg  Arg Leu Trp
    1050                1055                1060

ttt ttg  cac tgt ttg gaa gga  ccg agc ccg acc tat  aat att ccg       3734
Phe Leu  His Cys Leu Glu Gly  Pro Ser Pro Thr Tyr  Asn Ile Pro
    1065                1070                1075

gtt gtc  gtt cat tta act gga  gat ttg gac caa aag  gcg ctg gca       3779
Val Val  Val His Leu Thr Gly  Asp Leu Asp Gln Lys  Ala Leu Ala
    1080                1085                1090

gct gct  ctg ggt gat gtg gca  aca aga cat gag ccg  ctt cga aca       3824
Ala Ala  Leu Gly Asp Val Ala  Thr Arg His Glu Pro  Leu Arg Thr
    1095                1100                1105

att ttc  ccg gat cag caa ggg  aca aca cac cag ctg  ata ttg gaa       3869
Ile Phe  Pro Asp Gln Gln Gly  Thr Thr His Gln Leu  Ile Leu Glu
    1110                1115                1120

gag gat  caa tcg cgt ccc gaa  ctc act gta tcc cat  gtc agc gaa       3914
Glu Asp  Gln Ser Arg Pro Glu  Leu Thr Val Ser His  Val Ser Glu
    1125                1130                1135

cat gag  ttg gaa aaa gtc ctg  gct gag gcc gtc cgg  cac cgt tac       3959
His Glu  Leu Glu Lys Val Leu  Ala Glu Ala Val Arg  His Arg Tyr
    1140                1145                1150

cat tta  gaa aag gag cct ccg  ttt cgc gcc cag cta  ttc gta ctc       4004
His Leu  Glu Lys Glu Pro Pro  Phe Arg Ala Gln Leu  Phe Val Leu
    1155                1160                1165

gga ccg  gac aaa ttt gtg ctt  ctt ctg ctt ttg cac  cat atg atc       4049
```

Page 16

```
Gly Pro  Asp Lys Phe Val Leu  Leu Leu Leu Leu His  His Met Ile
    1170                 1175             1180

ggc gac  ggc tgg tct tta atg  ccg ctg acc cgt gat  ttg gaa acc        4094
Gly Asp  Gly Trp Ser Leu Met  Pro Leu Thr Arg Asp  Leu Glu Thr
    1185                 1190             1195

gcc tat  aat gcg cgt ctg caa  ggt gag gca cct gtt  tgg gag ccg        4139
Ala Tyr  Asn Ala Arg Leu Gln  Gly Glu Ala Pro Val  Trp Glu Pro
    1200                 1205             1210

ctt tct  ata caa tat gcc gac  tat gcc gta tgg cag  gaa tat ctg        4184
Leu Ser  Ile Gln Tyr Ala Asp  Tyr Ala Val Trp Gln  Glu Tyr Leu
    1215                 1220             1225

ctt ggc  agt gag aac aat ccg  gac agt ctg att gcc  aaa cag ctc        4229
Leu Gly  Ser Glu Asn Asn Pro  Asp Ser Leu Ile Ala  Lys Gln Leu
    1230                 1235             1240

gaa tat  tgg tcg aag gca ttg  gaa cat ctg cct gat  cag ctg gag        4274
Glu Tyr  Trp Ser Lys Ala Leu  Glu His Leu Pro Asp  Gln Leu Glu
    1245                 1250             1255

ctt ccg  act gat cac ccg cgt  ccg tcg gaa tct agc  tac aga agc        4319
Leu Pro  Thr Asp His Pro Arg  Pro Ser Glu Ser Ser  Tyr Arg Ser
    1260                 1265             1270

ggt acg  att gat ctc agc att  gac gaa cag ctg cac  ggc cgc ctg        4364
Gly Thr  Ile Asp Leu Ser Ile  Asp Glu Gln Leu His  Gly Arg Leu
    1275                 1280             1285

ttc gat  ttg tcc cgc agc acc  gga gtc agc atg ttt  atg att ttg        4409
Phe Asp  Leu Ser Arg Ser Thr  Gly Val Ser Met Phe  Met Ile Leu
    1290                 1295             1300

cag tcg  gct ctt gcc gca cta  ctg aca cgg ctc gga  gca ggc cat        4454
Gln Ser  Ala Leu Ala Ala Leu  Leu Thr Arg Leu Gly  Ala Gly His
    1305                 1310             1315

gat att  ccg ctc ggc agc ccg  ata gcg gga aga aac  gat gat gct        4499
Asp Ile  Pro Leu Gly Ser Pro  Ile Ala Gly Arg Asn  Asp Asp Ala
    1320                 1325             1330

tta ggc  gaa att gtc gga ttg  ttt gtt aat acg ctt  gtg ctt cga        4544
Leu Gly  Glu Ile Val Gly Leu  Phe Val Asn Thr Leu  Val Leu Arg
    1335                 1340             1345

aca gat  aca tcg ggt aat ccg  agt ttt cgg gag ctt  ctg aac aga        4589
Thr Asp  Thr Ser Gly Asn Pro  Ser Phe Arg Glu Leu  Leu Asn Arg
    1350                 1355             1360

gtc aga  aaa gta aac ctt gca  gcg tat gag cat cag  gac ctt ccg        4634
Val Arg  Lys Val Asn Leu Ala  Ala Tyr Glu His Gln  Asp Leu Pro
    1365                 1370             1375

ttc gag  cgt ctt gtt gaa gtg  ctt aat ccg cgc agg  tca aga gcg        4679
Phe Glu  Arg Leu Val Glu Val  Leu Asn Pro Arg Arg  Ser Arg Ala
    1380                 1385             1390

aga cat  ccg ctg ttc caa att  atg ctt gcc ttt caa  aat acg cct        4724
Arg His  Pro Leu Phe Gln Ile  Met Leu Ala Phe Gln  Asn Thr Pro
    1395                 1400             1405

gaa cct  gag ctt gac ctt tcc  ggc ctc aag tca gac  att gaa att        4769
Glu Pro  Glu Leu Asp Leu Ser  Gly Leu Lys Ser Asp  Ile Glu Ile
    1410                 1415             1420

cgc agc  gtc gga gca gcc aag  ttt gat ttg acg att  gag ctt cgc        4814
```

```
Arg Ser Val Gly Ala Ala Lys  Phe Asp Leu Thr Ile  Glu Leu Arg
    1425            1430               1435

gaa cat cga aaa gcg gac gga  aca ccc gca ggt atc  ggc gga ttt         4859
Glu His Arg Lys Ala Asp Gly  Thr Pro Ala Gly Ile  Gly Gly Phe
    1440            1445               1450

ctt gaa tac agc aca gat tta  ttt gag cga aat acg  gtt caa aca         4904
Leu Glu Tyr Ser Thr Asp Leu  Phe Glu Arg Asn Thr  Val Gln Thr
    1455            1460               1465

ttg gca gaa cgg ctt gtc agg  ctg ctg gat tca gca  gca gac gac         4949
Leu Ala Glu Arg Leu Val Arg  Leu Leu Asp Ser Ala  Ala Asp Asp
    1470            1475               1480

cct gat cag ccg atc gaa aaa  ctc gat att ctc ctt  ccg gcc gaa         4994
Pro Asp Gln Pro Ile Glu Lys  Leu Asp Ile Leu Leu  Pro Ala Glu
    1485            1490               1495

cgg gaa aac atg ctg gct gat  tgg agc aaa tcg tcc  aac agc ata         5039
Arg Glu Asn Met Leu Ala Asp  Trp Ser Lys Ser Ser  Asn Ser Ile
    1500            1505               1510

cct tgt tca tct ttg ccc gtt  ttg ttt gaa aag cag  gcg gca aag         5084
Pro Cys Ser Ser Leu Pro Val  Leu Phe Glu Lys Gln  Ala Ala Lys
    1515            1520               1525

gat ccg gaa gct gtt gcg gtt  ata tgt gaa aat aat  gcg ctc aca         5129
Asp Pro Glu Ala Val Ala Val  Ile Cys Glu Asn Asn  Ala Leu Thr
    1530            1535               1540

tac ggc gag ctt aat aaa cgc  gct aac cgg ctt gct  cat ctc ctg         5174
Tyr Gly Glu Leu Asn Lys Arg  Ala Asn Arg Leu Ala  His Leu Leu
    1545            1550               1555

atc gcc aaa ggg gtc ggc ccc  gaa caa ttt gct gcc  ctg gcg ctg         5219
Ile Ala Lys Gly Val Gly Pro  Glu Gln Phe Ala Ala  Leu Ala Leu
    1560            1565               1570

cca aga tct ttg gat atg gtt  gtc gga ttg ctt gca  gtg ctg aag         5264
Pro Arg Ser Leu Asp Met Val  Val Gly Leu Leu Ala  Val Leu Lys
    1575            1580               1585

gct ggc gcg gca tat gtg ccc  cta gat ccg gac tac  ccg gcg gaa         5309
Ala Gly Ala Ala Tyr Val Pro  Leu Asp Pro Asp Tyr  Pro Ala Glu
    1590            1595               1600

agg ata gcg ttt atg ttg aat  gat gcc cat cct gtc  tgc atc gta         5354
Arg Ile Ala Phe Met Leu Asn  Asp Ala His Pro Val  Cys Ile Val
    1605            1610               1615

acg agt tca gcg gta gaa tca  aat ctt tct gtt ccc  gga tct gtt         5399
Thr Ser Ser Ala Val Glu Ser  Asn Leu Ser Val Pro  Gly Ser Val
    1620            1625               1630

gaa agg att gta ctc gat gat  cct tgt ata caa gag  gaa ctg aaa         5444
Glu Arg Ile Val Leu Asp Asp  Pro Cys Ile Gln Glu  Glu Leu Lys
    1635            1640               1645

gga tgc gct gct gcg aat ccg  tgt gat gcc gac cgc  acc gcg cct         5489
Gly Cys Ala Ala Ala Asn Pro  Cys Asp Ala Asp Arg  Thr Ala Pro
    1650            1655               1660

ttg ctg cct ctt cac ccg gcg  tat gtc att tac aca  tcc ggt tcg         5534
Leu Leu Pro Leu His Pro Ala  Tyr Val Ile Tyr Thr  Ser Gly Ser
    1665            1670               1675

acc ggc aag ccg aaa gga gtc  gtt gtt ccg cat caa  aat gtc gtg         5579
```

Thr Gly Lys Pro Lys Gly Val Val Val Pro His Gln Asn Val Val
   1680              1685              1690

cgg ctg ttc gga gcg act gac caa tgg ttc cat ttc ggc gct gac    5624
Arg Leu Phe Gly Ala Thr Asp Gln Trp Phe His Phe Gly Ala Asp
   1695             1700            1705

gat gtt tgg acg atg ttt cat tct tac gca ttt gac ttt tcg gta    5669
Asp Val Trp Thr Met Phe His Ser Tyr Ala Phe Asp Phe Ser Val
   1710             1715            1720

tgg gaa att tgg gga gca ctt ctc aat ggg ggc cgg ctt atc gtg    5714
Trp Glu Ile Trp Gly Ala Leu Leu Asn Gly Gly Arg Leu Ile Val
   1725             1730            1735

gtg cct cat aca atc agc cgc tcg ccg gcc gaa ttt ctc aat ctt    5759
Val Pro His Thr Ile Ser Arg Ser Pro Ala Glu Phe Leu Asn Leu
   1740             1745            1750

ctc gtt cgc gag gga gtg act gtg ctc aat cag act cct tca gca    5804
Leu Val Arg Glu Gly Val Thr Val Leu Asn Gln Thr Pro Ser Ala
   1755             1760            1765

ttt tat caa ttg atg cag gct gac cgg gac aat gct gag aca gga    5849
Phe Tyr Gln Leu Met Gln Ala Asp Arg Asp Asn Ala Glu Thr Gly
   1770             1775            1780

aag ctc ctt tcc ttg cgc ttt atc att ttt gga ggg gaa gcg ctc    5894
Lys Leu Leu Ser Leu Arg Phe Ile Ile Phe Gly Gly Glu Ala Leu
   1785             1790            1795

gag ctg aag agg ctt gaa gat tgg tat gaa cgc cat ccg gat cat    5939
Glu Leu Lys Arg Leu Glu Asp Trp Tyr Glu Arg His Pro Asp His
   1800             1805            1810

ttc ccg cgg ctg att aat atg tat gga att aca gaa acg acg gtg    5984
Phe Pro Arg Leu Ile Asn Met Tyr Gly Ile Thr Glu Thr Thr Val
   1815             1820            1825

cat gtc agc tat att tct ctc gat cag cag act gcg gca ctt caa    6029
His Val Ser Tyr Ile Ser Leu Asp Gln Gln Thr Ala Ala Leu Gln
   1830             1835            1840

gcg aac agc ctg atc ggc cag ggg att ccc gat tta ggt gta tat    6074
Ala Asn Ser Leu Ile Gly Gln Gly Ile Pro Asp Leu Gly Val Tyr
   1845             1850            1855

gtg ctt gat gag tat ctt gaa ccg gtg ccc ccg ggc gtc acc gga    6119
Val Leu Asp Glu Tyr Leu Glu Pro Val Pro Pro Gly Val Thr Gly
   1860             1865            1870

gag atg tat gtg tcc ggg ggc ggt ctc gca cgg ggt tac ttg ggc    6164
Glu Met Tyr Val Ser Gly Gly Gly Leu Ala Arg Gly Tyr Leu Gly
   1875             1880            1885

aga ccc gat ttg aca gca gac cgc ttt gtc gct gat ccg ttt ggt    6209
Arg Pro Asp Leu Thr Ala Asp Arg Phe Val Ala Asp Pro Phe Gly
   1890             1895            1900

ccg ccg gga acc cgg atg tac cgc acc ggc gat ctt gcc cgc agg    6254
Pro Pro Gly Thr Arg Met Tyr Arg Thr Gly Asp Leu Ala Arg Arg
   1905             1910            1915

cgc cag gac ggc tct ctg gat tat atg ggg cgg gcc gac cag cag    6299
Arg Gln Asp Gly Ser Leu Asp Tyr Met Gly Arg Ala Asp Gln Gln
   1920             1925            1930

atc aag atc cgc ggc ttc aga atc gaa ctc ggc gag ata gaa gcg    6344

```
Ile Lys  Ile Arg Gly Phe Arg  Ile Glu Leu Gly Glu  Ile Glu Ala
    1935             1940             1945

gtg ctt  gtc cgc cat cat aga  gtg aat caa gcc gct  gtg gtc gta    6389
Val Leu  Val Arg His His Arg  Val Asn Gln Ala Ala  Val Val Val
    1950             1955             1960

agg gaa  ggc cag ccc ggg gac  aag cgt ctg atc gcc  tac gtt gtc    6434
Arg Glu  Gly Gln Pro Gly Asp  Lys Arg Leu Ile Ala  Tyr Val Val
    1965             1970             1975

ccg gcg  tct gaa gaa gaa acc  gat ccg gct gag ctg  cga cgg ttt    6479
Pro Ala  Ser Glu Glu Glu Thr  Asp Pro Ala Glu Leu  Arg Arg Phe
    1980             1985             1990

gcc gca  ggc act ctg ccg gag  tac atg gtg cct tcc  gca ttt gta    6524
Ala Ala  Gly Thr Leu Pro Glu  Tyr Met Val Pro Ser  Ala Phe Val
    1995             2000             2005

aaa att  tca gag ctg cca ttg  act cca aat ggc aaa  ctt gac cga    6569
Lys Ile  Ser Glu Leu Pro Leu  Thr Pro Asn Gly Lys  Leu Asp Arg
    2010             2015             2020

aaa gcg  ctg cct gag cct gac  ttt gcc gcg gct gtt  aaa ggg cgg    6614
Lys Ala  Leu Pro Glu Pro Asp  Phe Ala Ala Ala Val  Lys Gly Arg
    2025             2030             2035

ggc cca  aga aca ccg cag gaa  gaa atc ctg tgc gat  ctt ttt tca    6659
Gly Pro  Arg Thr Pro Gln Glu  Glu Ile Leu Cys Asp  Leu Phe Ser
    2040             2045             2050

gaa atc  ctc aat gcg ccg aga  gtc ggt att gat gac  gga ttc ttc    6704
Glu Ile  Leu Asn Ala Pro Arg  Val Gly Ile Asp Asp  Gly Phe Phe
    2055             2060             2065

gaa ctt  ggc ggg cat tct ctt  ctc gcc gta cag ctg  atg agc cgc    6749
Glu Leu  Gly Gly His Ser Leu  Leu Ala Val Gln Leu  Met Ser Arg
    2070             2075             2080

att cgt  gaa gcg ctc gga gtc  gaa ctc ggc atc ggc  gac ctg ctt    6794
Ile Arg  Glu Ala Leu Gly Val  Glu Leu Gly Ile Gly  Asp Leu Leu
    2085             2090             2095

gaa gcg  cct aca gtc agc gga  ctt gct gaa cgg ctt  gaa tcc ggc    6839
Glu Ala  Pro Thr Val Ser Gly  Leu Ala Glu Arg Leu  Glu Ser Gly
    2100             2105             2110

ggc agg  caa agt gcg ctg gat  gta atg ctt ccg ctg  aga acc gga    6884
Gly Arg  Gln Ser Ala Leu Asp  Val Met Leu Pro Leu  Arg Thr Gly
    2115             2120             2125

ggc agt  caa gat ccc ctg ttc  tgt gtg cac ccg gcc  gga ggc ctc    6929
Gly Ser  Gln Asp Pro Leu Phe  Cys Val His Pro Ala  Gly Gly Leu
    2130             2135             2140

agc tgg  tgt tat gcc ggt tta  atg acg gca ctc ggc  aag gaa tac    6974
Ser Trp  Cys Tyr Ala Gly Leu  Met Thr Ala Leu Gly  Lys Glu Tyr
    2145             2150             2155

ccg ata  tac ggc ctt caa gcc  cgc ggg atc gca agg  cag gaa gag    7019
Pro Ile  Tyr Gly Leu Gln Ala  Arg Gly Ile Ala Arg  Gln Glu Glu
    2160             2165             2170

cta cct  gat acg ctc gac gac  atg gct gct gat tat  att cgg cac    7064
Leu Pro  Asp Thr Leu Asp Asp  Met Ala Ala Asp Tyr  Ile Arg His
    2175             2180             2185

att cgt  acg att cag cca aca  ggc cct tat cgt ctg  ctc ggc tgg    7109
Ile Arg  Thr Ile Gln Pro Thr  Gly Pro Tyr Arg Leu  Leu Gly Trp
```

```
Ile Arg Thr Ile Gln Pro Thr  Gly Pro Tyr Arg Leu  Leu Gly Trp
2190              2195             2200
```

| | | | |
|---|---|---|---|
| tcc ctc gga ggc aat gtt gtt | cat gcg atc gca act | cag ctg cag | 7154 |
| Ser Leu Gly Gly Asn Val Val | His Ala Ile Ala Thr | Gln Leu Gln | |
| 2205 2210 | 2215 | | |

| | | | |
|---|---|---|---|
| gaa caa ggg gaa gac ata tcg | ctt ctt gtc atg ctt | gat gct tac | 7199 |
| Glu Gln Gly Glu Asp Ile Ser | Leu Leu Val Met Leu | Asp Ala Tyr | |
| 2220 2225 | 2230 | | |

| | | | |
|---|---|---|---|
| ccg aat cac ttc ctt cca att | aaa gat gcg cct gat | gag cag gaa | 7244 |
| Pro Asn His Phe Leu Pro Ile | Lys Asp Ala Pro Asp | Glu Gln Glu | |
| 2235 2240 | 2245 | | |

| | | | |
|---|---|---|---|
| gca ttg att gcc ctc ttg gca | ttg gga gga tac gac | ccg gac agc | 7289 |
| Ala Leu Ile Ala Leu Leu Ala | Leu Gly Gly Tyr Asp | Pro Asp Ser | |
| 2250 2255 | 2260 | | |

| | | | |
|---|---|---|---|
| ctt gac ggg gca ccg ctg aac | ctt tcc agt gcg atc | gat att ctg | 7334 |
| Leu Asp Gly Ala Pro Leu Asn | Leu Ser Ser Ala Ile | Asp Ile Leu | |
| 2265 2270 | 2275 | | |

| | | | |
|---|---|---|---|
| cgc cgc gac ggc agt gcg ctt | gcc agc ctg gac gag | gct gcg atc | 7379 |
| Arg Arg Asp Gly Ser Ala Leu | Ala Ser Leu Asp Glu | Ala Ala Ile | |
| 2280 2285 | 2290 | | |

| | | | |
|---|---|---|---|
| ttg aac tta aag gag acg tat | gtg aat tcc gtc cgg | att tta agc | 7424 |
| Leu Asn Leu Lys Glu Thr Tyr | Val Asn Ser Val Arg | Ile Leu Ser | |
| 2295 2300 | 2305 | | |

| | | | |
|---|---|---|---|
| gaa tat aag ccc cgc gtc ttt | cat ggc gat att ctg | ttt ttc aaa | 7469 |
| Glu Tyr Lys Pro Arg Val Phe | His Gly Asp Ile Leu | Phe Phe Lys | |
| 2310 2315 | 2320 | | |

| | | | |
|---|---|---|---|
| tca acc gtg ata ccg gaa tgg | ttt gat ccg atc gat | ccg gag tca | 7514 |
| Ser Thr Val Ile Pro Glu Trp | Phe Asp Pro Ile Asp | Pro Glu Ser | |
| 2325 2330 | 2335 | | |

| | | | |
|---|---|---|---|
| tgg cta cct tat ttg aac gga | aac att gat att cat | gat atg gat | 7559 |
| Trp Leu Pro Tyr Leu Asn Gly | Asn Ile Asp Ile His | Asp Met Asp | |
| 2340 2345 | 2350 | | |

| | | | |
|---|---|---|---|
| tgc cgg cat aaa gat ttg tgc | cag ccg gag ccg ctg | gcc gaa atc | 7604 |
| Cys Arg His Lys Asp Leu Cys | Gln Pro Glu Pro Leu | Ala Glu Ile | |
| 2355 2360 | 2365 | | |

| | | | |
|---|---|---|---|
| gga cgt cgg gtt tcg gaa aag | ctg gac gac ttg aaa | aaa gat acg | 7649 |
| Gly Arg Arg Val Ser Glu Lys | Leu Asp Asp Leu Lys | Lys Asp Thr | |
| 2370 2375 | 2380 | | |

| | | | |
|---|---|---|---|
| gat aag taggagaga tagcgatgac aaatcctttt gaaaatgaac acggcacata | | | 7705 |
| Asp Lys | | | |
| 2385 | | | |

| | |
|---|---|
| tgtcgtttta atgaatcagg atggccagta ttcactctgg ccggccttca ttcatattcc | 7765 |
| ggcagggtgg gaggtcgtct gcggggaagc gagccgcagc gcctgtatcg actatatcag | 7825 |
| ctcgaattgg accgatctct cgccgaacag catcaagctg gttgcagaag tacacggcgg | 7885 |
| tcagcaatga atagaaggcc gttcaatcaa aggacggttg tcagcattgt atatgtaaca | 7945 |
| gccatgttta tggctgcgat ggatgcgagt gtcatcaaca tggcattgcc tgccatcatg | 8005 |
| accgaatttc aaacattgcc gtctgctgca ggcacagtta acatcggata cttagtcagc | 8065 |
| ttggcggtct gtcttcccgc agccggctgg cttggggacc gctttgggac aaagcgtgtt | 8125 |

ttcctagccg cgctttcgat atttacagct gca                                    8158

<210>  8

<211>  2385

<212>  PRT

<213>  Bacillus licheniformis


<400>  8

Met Pro Glu Cys Gln His Asn Arg Lys Pro Leu Ser Gly Ala Gln Ala
1               5                   10                  15


Gly Ile Trp Phe Ala Gln Gln Leu Asp Pro Glu Asn Pro Ile Tyr Asn
            20                  25                  30


Thr Ala Glu Tyr Val Glu Ile Lys Gly Pro Leu Asp Gln Glu Leu Phe
        35                  40                  45


Glu Lys Ala Leu Arg His Val Ile Lys Glu Ala Glu Ser Phe His Ala
    50                  55                  60


Arg Phe Gly Glu Asp Gln Asp Gly Pro Trp Gln Glu Ile Val Pro Ser
65                  70                  75                  80


Thr Asp Phe Pro Leu His Tyr Ile Asp Val Ser Ser Glu Thr Asp Pro
                85                  90                  95


Glu Gln Ala Ala Lys Ser Trp Met Met Asp Asp Leu Ala Arg Pro Val
                100                 105                 110


Asp Leu Thr Arg Gly Pro Leu Phe Thr Glu Ala Leu Phe Lys Ala Ala
        115                 120                 125


Gln Asp His Tyr Phe Trp Tyr Gln Arg Thr His His Ile Ala Thr Asp
    130                 135                 140


Gly Phe Ser Phe Thr Leu Ile Ala Glu Arg Leu Ser Lys Ile Tyr Thr
145                 150                 155                 160


Ala Leu Met Gln Asn Lys Ser Ile Asp Gln Ser Gly Ala Phe Gly Ser
                165                 170                 175


Leu Asp Leu Ile Leu Ala Glu Glu Thr Ala Tyr Arg Ala Ser Glu Gln
                180                 185                 190


Tyr Gln Glu Asp Arg Arg Phe Trp Leu Gly Arg Phe Ser Asp Glu Pro
            195                 200                 205


**Page 22**

```
Glu Val Val Ser Leu Ala Glu Arg Ala Pro Arg Thr Ser Ser Ser Phe
210             215             220

Leu Arg Arg Ser Glu His Leu Pro Ser Glu Asp Ala Asp Arg Leu Leu
225             230             235             240

Ser Ala Ala Ser Arg Met Gly Ala Thr Trp His Glu Thr Val Met Ala
            245             250             255

Ala Ala Ala Ile Tyr Val His Arg Leu Thr Gly Ala Asn Asp Val Val
            260             265             270

Leu Gly Met Pro Met Met Cys Arg Leu Gly Ser Ala Ala Leu His Ile
        275             280             285

Pro Gly Met Val Met Asn Leu Leu Pro Leu Arg Ile Gly Met Gln Pro
    290             295             300

Gln Met Ser Ile Gly Glu Leu Val Lys Gln Ile Ser Gly Glu Met Met
305             310             315             320

Lys Leu Arg Arg His Gln His Tyr Arg His Glu Glu Leu Arg Arg Asp
            325             330             335

Leu Lys Leu Leu Gly Glu Asn Gln Arg Leu Phe Gly Pro Gln Leu Asn
        340             345             350

Leu Met Pro Phe Glu Asn Arg Leu Asn Phe Ala Gly Cys Gln Gly Ile
        355             360             365

Val His Asn Leu Ala Thr Gly Pro Val Asp Asp Leu Ser Ile Asn Ile
    370             375             380

Tyr Gly Arg Pro Asp Gly Gly Gly Leu Lys Val Asn Met Asp Ala Asn
385             390             395             400

Pro Ala Val Tyr His Ala Asp Glu Leu Glu Asp His Gly Asn Arg Phe
            405             410             415

Leu Thr Leu Leu Lys Thr Ile Ala Val Cys Glu Gln Thr Gln Pro Val
        420             425             430

Gly Thr Leu Asp Ile Leu Leu Pro Glu Glu Arg Ala Gln Ile Leu Ile
        435             440             445

Glu Trp Asn Gln Thr Glu His Gly Leu Pro Lys Glu Ser Leu Pro Gln
    450             455             460

Arg Phe Glu Arg Met Ala Lys Glu Cys Pro Glu Ser Pro Ala Val Val
465             470             475             480
```

Page 23

```
Cys Asn Asp Lys Val Leu Thr Tyr Ser Glu Leu Asn Gln Lys Ala Asn
            485                 490                 495

Gln Leu Ala His Leu Leu Ile Asp Gln Gly Ala Lys Pro Glu Thr Phe
            500                 505                 510

Ile Ala Leu Ala Leu Pro Arg Ser Ala Glu Met Val Val Ser Met Leu
            515                 520                 525

Ala Val Leu Lys Ala Gly Ala Ala Tyr Leu Pro Ile Asp Pro Asp Tyr
        530                 535                 540

Pro Ala Asp Arg Ile Glu Tyr Met Leu Asn Asp Ala Gln Pro Leu Leu
545                 550                 555                 560

Val Met Thr Ser Lys Glu Ala Gln Asp Thr Ile Gly Ser Gln Met Pro
                565                 570                 575

Lys Leu Ile Leu Asp Glu Gln Thr Val Met Glu Lys Met Ser Gly Cys
            580                 585                 590

Ser Glu Glu Asn Pro Gly Glu Gln His Ser Gly Asn Gln Pro Ala Tyr
        595                 600                 605

Met Ile Tyr Thr Ser Gly Ser Thr Gly Arg Pro Lys Gly Val Val Val
    610                 615                 620

Gln Ala Glu Ser Leu Phe Asn Phe Leu Leu Ser Met Gln Asp Met Phe
625                 630                 635                 640

Ala Leu Asn Gln Asp Asp Arg Leu Leu Ala Val Thr Thr Ile Ala Phe
                645                 650                 655

Asp Ile Ser Ala Leu Glu Ile Tyr Leu Pro Leu Ile Ser Gly Ser Ala
            660                 665                 670

Val Val Leu Ala Glu Lys Glu Thr Val Gln Asp Pro Ser Glu Leu Ala
        675                 680                 685

Lys Met Ile Glu Thr Tyr Glu Ile Thr Ile Met Gln Ala Thr Pro Thr
    690                 695                 700

Leu Trp His Ala Leu Ala Ser Ser Ala Pro Glu Lys Leu Lys Gly Leu
705                 710                 715                 720

Arg Ala Leu Val Gly Gly Glu Ala Leu Gln Ser Ser Leu Ala Arg Gln
                725                 730                 735

Leu Gln Gln Leu Ala Cys Ser Leu Thr Asn Leu Tyr Gly Pro Thr Glu
            740                 745                 750
```

Page 24

```
Thr Thr Ile Trp Ser Thr Ala Ala Ala Leu Glu Gly Asn Cys Thr Glu
        755             760             765

Pro Pro Gly Ile Gly Cys Ala Ile Trp Asn Thr Gln Leu Tyr Val Leu
    770             775             780

Asp Ala Gly Leu Gln Pro Val Pro Pro Gly Thr Ala Gly Glu Leu Tyr
785             790             795             800

Val Ala Gly Thr Gly Val Ala Arg Gly Tyr Leu Asn Arg His Ala Leu
            805             810             815

Thr Ala Glu Arg Phe Ile Ala Asn Pro Tyr Gly Pro Pro Gly Ser Arg
        820             825             830

Met Tyr Arg Thr Gly Asp Ile Val Arg Trp Arg Glu Asp Gly Ser Leu
        835             840             845

Asp Tyr Ile Gly Arg Ala Asp His Gln Val Lys Ile Arg Gly Phe Arg
    850             855             860

Ile Glu Met Gly Glu Ile Glu Ala Val Leu Ala Asn His Pro Val Val
865             870             875             880

Lys Gln Ala Ala Ala Ile Val Arg Glu Asp Gln Pro Gly Asp Pro Arg
            885             890             895

Leu Phe Ala Tyr Ala Val Pro Ala Ser Gly Glu Ser Leu Asp Pro Ala
        900             905             910

Glu Leu Arg Arg Phe Val Gly Glu Thr Leu Pro Asp Tyr Met Ile Pro
        915             920             925

Ser Ala Phe Val Ile Leu Asp Glu Leu Pro Leu Thr Pro Asn Gly Lys
    930             935             940

Leu Asp Arg Lys Ser Leu Pro Ala Pro Ala Val Ser Met His Thr Gly
945             950             955             960

Gly Arg Glu Pro Arg Thr Pro Gln Glu Glu Ile Leu Cys Asp Leu Phe
            965             970             975

Ala Glu Val Leu Gly Val Pro Arg Val Ser Ile Asp Asp Ser Phe Phe
            980             985             990

Asp Leu Gly Gly His Ser Leu Leu Ala Gly Arg Leu Val Gly Arg Ile
    995             1000            1005

Arg Glu Met Leu Gly Val Glu Leu Gly Ile Gly Arg Leu Phe Asp
    1010            1015            1020
```

**Page 25**

Glu Pro Thr Ala Ala Gly Leu Ala Lys Gln Leu Asp Gln Ala Gln
    1025              1030              1035

Ser Ala Arg Pro Ala Leu Arg Lys Arg Glu Arg Arg Lys Glu Ile
    1040              1045              1050

Pro Leu Ser Phe Ala Gln Arg Arg Leu Trp Phe Leu His Cys Leu
    1055              1060              1065

Glu Gly Pro Ser Pro Thr Tyr Asn Ile Pro Val Val Val His Leu
    1070              1075              1080

Thr Gly Asp Leu Asp Gln Lys Ala Leu Ala Ala Ala Leu Gly Asp
    1085              1090              1095

Val Ala Thr Arg His Glu Pro Leu Arg Thr Ile Phe Pro Asp Gln
    1100              1105              1110

Gln Gly Thr Thr His Gln Leu Ile Leu Glu Glu Asp Gln Ser Arg
    1115              1120              1125

Pro Glu Leu Thr Val Ser His Val Ser Glu His Glu Leu Glu Lys
    1130              1135              1140

Val Leu Ala Glu Ala Val Arg His Arg Tyr His Leu Glu Lys Glu
    1145              1150              1155

Pro Pro Phe Arg Ala Gln Leu Phe Val Leu Gly Pro Asp Lys Phe
    1160              1165              1170

Val Leu Leu Leu Leu Leu His His Met Ile Gly Asp Gly Trp Ser
    1175              1180              1185

Leu Met Pro Leu Thr Arg Asp Leu Glu Thr Ala Tyr Asn Ala Arg
    1190              1195              1200

Leu Gln Gly Glu Ala Pro Val Trp Glu Pro Leu Ser Ile Gln Tyr
    1205              1210              1215

Ala Asp Tyr Ala Val Trp Gln Glu Tyr Leu Leu Gly Ser Glu Asn
    1220              1225              1230

Asn Pro Asp Ser Leu Ile Ala Lys Gln Leu Glu Tyr Trp Ser Lys
    1235              1240              1245

Ala Leu Glu His Leu Pro Asp Gln Leu Glu Leu Pro Thr Asp His
    1250              1255              1260

Pro Arg Pro Ser Glu Ser Ser Tyr Arg Ser Gly Thr Ile Asp Leu
    1265              1270              1275

**Page 26**

Ser Ile Asp Glu Gln Leu His Gly Arg Leu Phe Asp Leu Ser Arg
    1280                    1285                1290

Ser Thr Gly Val Ser Met Phe Met Ile Leu Gln Ser Ala Leu Ala
    1295                    1300                1305

Ala Leu Leu Thr Arg Leu Gly Ala Gly His Asp Ile Pro Leu Gly
    1310                    1315                1320

Ser Pro Ile Ala Gly Arg Asn Asp Asp Ala Leu Gly Glu Ile Val
    1325                    1330                1335

Gly Leu Phe Val Asn Thr Leu Val Leu Arg Thr Asp Thr Ser Gly
    1340                    1345                1350

Asn Pro Ser Phe Arg Glu Leu Leu Asn Arg Val Arg Lys Val Asn
    1355                    1360                1365

Leu Ala Ala Tyr Glu His Gln Asp Leu Pro Phe Glu Arg Leu Val
    1370                    1375                1380

Glu Val Leu Asn Pro Arg Arg Ser Arg Ala Arg His Pro Leu Phe
    1385                    1390                1395

Gln Ile Met Leu Ala Phe Gln Asn Thr Pro Glu Pro Glu Leu Asp
    1400                    1405                1410

Leu Ser Gly Leu Lys Ser Asp Ile Glu Ile Arg Ser Val Gly Ala
    1415                    1420                1425

Ala Lys Phe Asp Leu Thr Ile Glu Leu Arg Glu His Arg Lys Ala
    1430                    1435                1440

Asp Gly Thr Pro Ala Gly Ile Gly Gly Phe Leu Glu Tyr Ser Thr
    1445                    1450                1455

Asp Leu Phe Glu Arg Asn Thr Val Gln Thr Leu Ala Glu Arg Leu
    1460                    1465                1470

Val Arg Leu Leu Asp Ser Ala Ala Asp Asp Pro Asp Gln Pro Ile
    1475                    1480                1485

Glu Lys Leu Asp Ile Leu Leu Pro Ala Glu Arg Glu Asn Met Leu
    1490                    1495                1500

Ala Asp Trp Ser Lys Ser Ser Asn Ser Ile Pro Cys Ser Ser Leu
    1505                    1510                1515

Pro Val Leu Phe Glu Lys Gln Ala Ala Lys Asp Pro Glu Ala Val
    1520                    1525                1530

Page 27

Ala Val Ile Cys Glu Asn Asn Ala Leu Thr Tyr Gly Glu Leu Asn
    1535              1540              1545

Lys Arg Ala Asn Arg Leu Ala His Leu Leu Ile Ala Lys Gly Val
    1550              1555              1560

Gly Pro Glu Gln Phe Ala Ala Leu Ala Leu Pro Arg Ser Leu Asp
    1565              1570              1575

Met Val Val Gly Leu Leu Ala Val Leu Lys Ala Gly Ala Ala Tyr
    1580              1585              1590

Val Pro Leu Asp Pro Asp Tyr Pro Ala Glu Arg Ile Ala Phe Met
    1595              1600              1605

Leu Asn Asp Ala His Pro Val Cys Ile Val Thr Ser Ser Ala Val
    1610              1615              1620

Glu Ser Asn Leu Ser Val Pro Gly Ser Val Glu Arg Ile Val Leu
    1625              1630              1635

Asp Asp Pro Cys Ile Gln Glu Glu Leu Lys Gly Cys Ala Ala Ala
    1640              1645              1650

Asn Pro Cys Asp Ala Asp Arg Thr Ala Pro Leu Leu Pro Leu His
    1655              1660              1665

Pro Ala Tyr Val Ile Tyr Thr Ser Gly Ser Thr Gly Lys Pro Lys
    1670              1675              1680

Gly Val Val Val Pro His Gln Asn Val Val Arg Leu Phe Gly Ala
    1685              1690              1695

Thr Asp Gln Trp Phe His Phe Gly Ala Asp Asp Val Trp Thr Met
    1700              1705              1710

Phe His Ser Tyr Ala Phe Asp Phe Ser Val Trp Glu Ile Trp Gly
    1715              1720              1725

Ala Leu Leu Asn Gly Gly Arg Leu Ile Val Val Pro His Thr Ile
    1730              1735              1740

Ser Arg Ser Pro Ala Glu Phe Leu Asn Leu Leu Val Arg Glu Gly
    1745              1750              1755

Val Thr Val Leu Asn Gln Thr Pro Ser Ala Phe Tyr Gln Leu Met
    1760              1765              1770

Gln Ala Asp Arg Asp Asn Ala Glu Thr Gly Lys Leu Leu Ser Leu
    1775              1780              1785

Page 28      ·

Arg Phe  Ile Ile Phe Gly Gly  Glu Ala Leu Glu Leu  Lys Arg Leu
    1790                1795                1800

Glu Asp  Trp Tyr Glu Arg His  Pro Asp His Phe Pro  Arg Leu Ile
    1805                1810                1815

Asn Met  Tyr Gly Ile Thr Glu  Thr Thr Val His Val  Ser Tyr Ile
    1820                1825                1830

Ser Leu  Asp Gln Gln Thr Ala  Ala Leu Gln Ala Asn  Ser Leu Ile
    1835                1840                1845

Gly Gln  Gly Ile Pro Asp Leu  Gly Val Tyr Val Leu  Asp Glu Tyr
    1850                1855                1860

Leu Glu  Pro Val Pro Pro Gly  Val Thr Gly Glu Met  Tyr Val Ser
    1865                1870                1875

Gly Gly  Gly Leu Ala Arg Gly  Tyr Leu Gly Arg Pro  Asp Leu Thr
    1880                1885                1890

Ala Asp  Arg Phe Val Ala Asp  Pro Phe Gly Pro Pro  Gly Thr Arg
    1895                1900                1905

Met Tyr  Arg Thr Gly Asp Leu  Ala Arg Arg Arg Gln  Asp Gly Ser
    1910                1915                1920

Leu Asp  Tyr Met Gly Arg Ala  Asp Gln Gln Ile Lys  Ile Arg Gly
    1925                1930                1935

Phe Arg  Ile Glu Leu Gly Glu  Ile Glu Ala Val Leu  Val Arg His
    1940                1945                1950

His Arg  Val Asn Gln Ala Ala  Val Val Val Arg Glu  Gly Gln Pro
    1955                1960                1965

Gly Asp  Lys Arg Leu Ile Ala  Tyr Val Val Pro Ala  Ser Glu Glu
    1970                1975                1980

Glu Thr  Asp Pro Ala Glu Leu  Arg Arg Phe Ala Ala  Gly Thr Leu
    1985                1990                1995

Pro Glu  Tyr Met Val Pro Ser  Ala Phe Val Lys Ile  Ser Glu Leu
    2000                2005                2010

Pro Leu  Thr Pro Asn Gly Lys  Leu Asp Arg Lys Ala  Leu Pro Glu
    2015                2020                2025

Pro Asp  Phe Ala Ala Ala Val  Lys Gly Arg Gly Pro  Arg Thr Pro
    2030                2035                2040

**Page 29**

```
Gln Glu  Glu Ile Leu Cys Asp  Leu Phe Ser Glu Ile  Leu Asn Ala
    2045             2050             2055

Pro Arg  Val Gly Ile Asp Asp  Gly Phe Phe Glu Leu  Gly Gly His
    2060             2065             2070

Ser Leu  Leu Ala Val Gln Leu  Met Ser Arg Ile Arg  Glu Ala Leu
    2075             2080             2085

Gly Val  Glu Leu Gly Ile Gly  Asp Leu Leu Glu Ala  Pro Thr Val
    2090             2095             2100

Ser Gly  Leu Ala Glu Arg Leu  Glu Ser Gly Gly Arg  Gln Ser Ala
    2105             2110             2115

Leu Asp  Val Met Leu Pro Leu  Arg Thr Gly Gly Ser  Gln Asp Pro
    2120             2125             2130

Leu Phe  Cys Val His Pro Ala  Gly Gly Leu Ser Trp  Cys Tyr Ala
    2135             2140             2145

Gly Leu  Met Thr Ala Leu Gly  Lys Glu Tyr Pro Ile  Tyr Gly Leu
    2150             2155             2160

Gln Ala  Arg Gly Ile Ala Arg  Gln Glu Glu Leu Pro  Asp Thr Leu
    2165             2170             2175

Asp Asp  Met Ala Ala Asp Tyr  Ile Arg His Ile Arg  Thr Ile Gln
    2180             2185             2190

Pro Thr  Gly Pro Tyr Arg Leu  Leu Gly Trp Ser Leu  Gly Gly Asn
    2195             2200             2205

Val Val  His Ala Ile Ala Thr  Gln Leu Gln Glu Gln  Gly Glu Asp
    2210             2215             2220

Ile Ser  Leu Leu Val Met Leu  Asp Ala Tyr Pro Asn  His Phe Leu
    2225             2230             2235

Pro Ile  Lys Asp Ala Pro Asp  Glu Gln Glu Ala Leu  Ile Ala Leu
    2240             2245             2250

Leu Ala  Leu Gly Gly Tyr Asp  Pro Asp Ser Leu Asp  Gly Ala Pro
    2255             2260             2265

Leu Asn  Leu Ser Ser Ala Ile  Asp Ile Leu Arg Arg  Asp Gly Ser
    2270             2275             2280

Ala Leu  Ala Ser Leu Asp Glu  Ala Ala Ile Leu Asn  Leu Lys Glu
    2285             2290             2295
```

Page 30

```
Thr Tyr Val Asn Ser Val Arg Ile Leu Ser Glu Tyr Lys Pro Arg
    2300                2305            2310

Val Phe His Gly Asp Ile Leu Phe Phe Lys Ser Thr Val Ile Pro
    2315                2320            2325

Glu Trp Phe Asp Pro Ile Asp Pro Glu Ser Trp Leu Pro Tyr Leu
    2330                2335            2340

Asn Gly Asn Ile Asp Ile His Asp Met Asp Cys Arg His Lys Asp
    2345                2350            2355

Leu Cys Gln Pro Glu Pro Leu Ala Glu Ile Gly Arg Arg Val Ser
    2360                2365            2370

Glu Lys Leu Asp Asp Leu Lys Lys Asp Thr Asp Lys
    2375                2380            2385
```

<210> 9

<211> 1000

<212> DNA

<213> Bacillus licheniformis


<220>

<221> CDS

<222> (321)..(998)

<223>


<400> 9

```
ttccgatcca tatcatgaaa gtgaagggtg ggataaatgt atactctact tcgaacacgc      60

ttcggaagct aaagcagagc ggaaacgaag tccgatcttt aaaggcgtac gaacgcagtc     120

gatcttaagc gtcatacgca aactccgtcg cgactaggtt cggaccgcag caatgagccc     180

tggaaggtag ctgatgcttc aatttggccg gttgtccgca gctgatgggg tatgcatgag     240

tcagcggggg ttcaaaagtc tgaggatgcc tttatttaaa tgtgtacgca cccgaagagg     300

cggacgggga tctgcctgtt atg gtg tgg att cat ggg ggc gct ttt tat cgc    353
                        Met Val Trp Ile His Gly Gly Ala Phe Tyr Arg
                          1               5                  10

ggc gcc gga agt gaa ccg ctc tat gac ggg act cag ctt gca aag cag      401
Gly Ala Gly Ser Glu Pro Leu Tyr Asp Gly Thr Gln Leu Ala Lys Gln
              15                  20                  25

gga aag gtg atc gtg gtc acc atc aat tat cgc ctc ggt ccg ttc ggt      449
Gly Lys Val Ile Val Val Thr Ile Asn Tyr Arg Leu Gly Pro Phe Gly
          30                  35                  40

ttt ttg cat cta tcc tca att gat gat tcc tac agc agc aat ctt ggc      497
```

Page 31

```
Phe Leu His Leu Ser Ser Ile Asp Asp Ser Tyr Ser Ser Asn Leu Gly
    45                  50              55

ctg ctg gat caa atc gcg gct ctc gag tgg gtg aaa gac aat atc gct    545
Leu Leu Asp Gln Ile Ala Ala Leu Glu Trp Val Lys Asp Asn Ile Ala
60              65              70                      75

ttc ttt ggc gga gac cgt cat cac att acg gtt ttt gga gag tcg gcg    593
Phe Phe Gly Gly Asp Arg His His Ile Thr Val Phe Gly Glu Ser Ala
            80              85              90

gga tcg atg agc atc gct tcg ctt ttg gcg atg ccg aaa gca aag ggg    641
Gly Ser Met Ser Ile Ala Ser Leu Leu Ala Met Pro Lys Ala Lys Gly
            95              100             105

ctt ttt caa cag gcc att atg gaa agc ggg gct tcc gca act atg tcc    689
Leu Phe Gln Gln Ala Ile Met Glu Ser Gly Ala Ser Ala Thr Met Ser
        110             115             120

gat aag ctt gcg aaa gct gca gca gaa aga ttc tta agg att ctc gat    737
Asp Lys Leu Ala Lys Ala Ala Ala Glu Arg Phe Leu Arg Ile Leu Asp
        125             130             135

att gat cat cat cat ctg gag cgc ctt cat gat gta tct gat caa gaa    785
Ile Asp His His His Leu Glu Arg Leu His Asp Val Ser Asp Gln Glu
140             145             150             155

ctt ctt gaa gcc gcc gat cag ctg cgc act tta atg gga gaa aat att    833
Leu Leu Glu Ala Ala Asp Gln Leu Arg Thr Leu Met Gly Glu Asn Ile
            160             165             170

ttt gaa ttg att ttt ctg cct gcg ctt gac gaa aaa acc ttg ccg ctg    881
Phe Glu Leu Ile Phe Leu Pro Ala Leu Asp Glu Lys Thr Leu Pro Leu
            175             180             185

aag ccg gag gtc gcc gtc gca aaa ggc gcg gca aaa gag atc aat cta    929
Lys Pro Glu Val Ala Val Ala Lys Gly Ala Ala Lys Glu Ile Asn Leu
        190             195             200

tta atc gga acc aaa ccc gtg atg aag gcg tct gtt ttt tcc tct gat    977
Leu Ile Gly Thr Lys Pro Val Met Lys Ala Ser Val Phe Ser Ser Asp
        205             210             215

tcg aaa ctt gct gag agc aat ag                                    1000
Ser Lys Leu Ala Glu Ser Asn
220                 225
```

<210> 10

<211> 226

<212> PRT

<213> Bacillus licheniformis

<400> 10

```
Met Val Trp Ile His Gly Gly Ala Phe Tyr Arg Gly Ala Gly Ser Glu
1               5               10                      15

Pro Leu Tyr Asp Gly Thr Gln Leu Ala Lys Gln Gly Lys Val Ile Val
            20              25              30
```

```
Val Thr Ile Asn Tyr Arg Leu Gly Pro Phe Gly Phe Leu His Leu Ser
        35                  40                  45

Ser Ile Asp Asp Ser Tyr Ser Ser Asn Leu Gly Leu Leu Asp Gln Ile
        50              55                  60

Ala Ala Leu Glu Trp Val Lys Asp Asn Ile Ala Phe Phe Gly Gly Asp
65                  70                  75                  80

Arg His His Ile Thr Val Phe Gly Glu Ser Ala Gly Ser Met Ser Ile
                85                  90                  95

Ala Ser Leu Leu Ala Met Pro Lys Ala Lys Gly Leu Phe Gln Gln Ala
            100                 105                 110

Ile Met Glu Ser Gly Ala Ser Ala Thr Met Ser Asp Lys Leu Ala Lys
        115                 120                 125

Ala Ala Ala Glu Arg Phe Leu Arg Ile Leu Asp Ile Asp His His His
    130                 135                 140

Leu Glu Arg Leu His Asp Val Ser Asp Gln Glu Leu Leu Glu Ala Ala
145                 150                 155                 160

Asp Gln Leu Arg Thr Leu Met Gly Glu Asn Ile Phe Glu Leu Ile Phe
            165                 170                 175

Leu Pro Ala Leu Asp Glu Lys Thr Leu Pro Leu Lys Pro Glu Val Ala
            180                 185                 190

Val Ala Lys Gly Ala Ala Lys Glu Ile Asn Leu Leu Ile Gly Thr Lys
        195                 200                 205

Pro Val Met Lys Ala Ser Val Phe Ser Ser Asp Ser Lys Leu Ala Glu
210                 215                 220

Ser Asn
225
```